

Study Title

An F₁-Extended One-Generation Reproduction Toxicity Study in Crl:CD(SD) Rats with
2,4-Dichlorophenoxyacetic Acid

Data Requirement

Non-Data Requirement Report

Authors

Mary Sue Marty, Barbara H. Neal, Carol L. Zablotny, Barry L. Yano, Amanda K. Andrus,
Michael R. Woolhiser, Darrell R. Boverhof, Shakil A. Saghir, Adam W. Perala, Julie K.
Passage, Marie A. Lawson, James S. Bus, James C. Lamb IV and Larry Hammond

Study Completed on:

September 5, 2013

Performing Laboratory

Toxicology and Environmental Research and Consulting
The Dow Chemical Company
Midland MI
and
Center for Toxicology and Mechanistic Biology,
Exponent, Inc.
Alexandria VA
and
Intrinsic Environmental Sciences Inc.
Mississauga, Ontario, Canada
and
Department of Biological and Biomedical Sciences
Aga Khan University
Karachi, Pakistan
and
Industry Task Force II on 2,4-D Research Data
Washington, DC

Published by:

Toxicological Sciences

Publication Reference ID

Toxicological Sciences
10.1093/toxsci/kft213

STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS

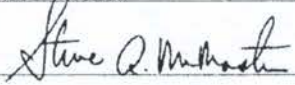
TITLE: An F₁-Extended One-Generation Reproductive Toxicity Study in
Crl:CD(SD) Rats with 2,4-Dichlorophenoxyacetic Acid

No claim of confidentiality is made for any information contained in this study on the basis of its falling within the scope of FIFRA Section 10(d)(1)(A),(B) or (C).

Company: Industry Task Force II on 2,4-D Research Data

Company Agent: Steve A. McMaster

Title: Technical Director

Signature: 

Date: 24 March 2017



**STATEMENT OF COMPLIANCE WITH
GOOD LABORATORY PRACTICE STANDARDS**

This study is not subject to Good Laboratory Practice Standards 40 CFR Part 160.

Submitter:

Company: Industry Task Force II on 2,4-D Research Data

Company Agent: Steve A. McMaster

Title: Technical Director

Signature: 

Date: 24 March 2017



An F₁-Extended One-Generation Reproductive Toxicity Study in Crl:CD(SD) Rats With 2,4-Dichlorophenoxyacetic Acid

Mary Sue Marty,^{*,†} Barbara H. Neal,[‡] Carol L. Zablotny,^{*} Barry L. Yano,^{*} Amanda K. Andrus,^{*} Michael R. Woolhiser,^{*} Darrell R. Boverhof,^{*} Shakil A. Saghir,^{‡,§} Adam W. Perala,^{*} Julie K. Passage,^{*} Marie A. Lawson,^{*} James S. Bus,[†] James C. Lamb IV,[†] and Larry Hammond[¶]

^{*}Toxicology and Environmental Research & Consulting, The Dow Chemical Company, Midland, Michigan; [†]Center for Toxicology and Mechanistic Biology, Exponent, Inc., Alexandria, Virginia; [‡]Intrinsic Environmental Sciences Inc., Mississauga, Ontario, Canada; [§]Department of Biological and Biomedical Sciences, Aga Khan University, Karachi, Pakistan; and [¶]Industry Task Force II on 2,4-D Research Data, Washington, District of Columbia

[†]To whom correspondence should be addressed at Toxicology and Environmental Research & Consulting, The Dow Chemical Company, 1803 Building, Midland, MI 48674. Fax: (517) 638-9863. E-mail: bneal@exponent.com.

Received June 13, 2013; accepted September 5, 2013

2,4-Dichlorophenoxyacetic acid (2,4-D) was assessed for systemic toxicity, reproductive toxicity, developmental neurotoxicity (DNT), developmental immunotoxicity (DIT), and endocrine toxicity. CD rats (27/sex/dose) were exposed to 0, 100, 300, 600 (female), or 800 (male) ppm 2,4-D in diet. Nonlinear toxicokinetic behavior was shown at high doses; the renal clearance saturation threshold for 2,4-D was exceeded markedly in females and slightly exceeded in males. Exposure was 4 weeks pre mating, 7 weeks post mating for P1 males and through lactation for P1 females. F1 offspring were examined for survival and development, and at weaning, pups were divided in cohorts, by sex and dose, and by systemic toxicity (10), DNT (10), DIT (20), and reproductive toxicity (≥ 23). Remaining weanlings were evaluated for systemic toxicity and neuropathology (10–12). Body weight decreased during lactation in high-dose P1 females and in F1 pups. Kidney was the primary target organ, with slight degeneration of proximal convoluted tubules observed in high-dose P1 males and in high-dose F1 males and females. A slight intergenerational difference in kidney toxicity was attributed to increased intake of 2,4-D in F1 offspring. Decreased weanling testes weights and delayed preputial separation in F1 males were attributed to decreased body weights. Endocrine-related effects were limited to slight thyroid hormone changes and adaptive histopathology in high-dose GD 17 dams seen only at a nonlinear toxicokinetic dose. 2,4-D did not cause reproductive toxicity, DNT, or DIT. The “No Observed Adverse Effect Level” for systemic toxicity was 300 ppm in both males (16.6 mg/kg/day) and females (20.6 mg/kg/day), which is approximately 6700- to 93 000-fold higher than that reported for 2,4-D exposures in human biomonitoring studies.

Key Words: 2,4-D, extended one-generation reproductive toxicity study; EOGRTS; endocrine; androgen; estrogen; thyroid; developmental immunotoxicity; developmental neurotoxicity; reproductive toxicity; systemic toxicity; kinetically derived maximum dose; KMD; toxicokinetics.

The Industry Task Force II on 2,4-D Research Data conducted a newly designed Extended One-Generation Reproductive Toxicity Study (EOGRTS) to supplement the 2,4-dichlorophenoxyacetic acid (2,4-D) toxicology data set. This EOGRTS was a modification of the study design proposed by Cooper *et al.* (2006) and preceded the recently published OECD 443 test guideline (OECD, 2011). The study was intended to provide a thorough toxicity assessment across life stages by a combined examination of reproductive toxicity, developmental neurotoxicity (DNT), developmental immunotoxicity (DIT), endocrine disruption (including thyroid parameters), and systemic toxicity. The EOGRTS is one component of the pesticide toxicology test battery advanced by the International Life Sciences Institute–Health and Environmental Sciences Institute Agricultural Chemicals Safety Assessment (ACSA) workgroup (Doe *et al.*, 2006) and advocates the use of toxicokinetic data for dose selection to avoid “high doses that are irrelevant due to saturation of metabolic processes” (Carmichael *et al.*, 2006). In addition, the ACSA EOGRTS protocol was intended to significantly reduce use of test animals by combining multiple toxicity endpoints commonly evaluated in independent tests into a single protocol (reproduction, DNT, DIT, and endocrine) and by potentially eliminating the requirement to conduct a second-generation mating to assess reproduction (Fegert *et al.*, 2012).

The toxicity of 2,4-D has previously been characterized in a comprehensive series of regulatory-mandated and other toxicity tests (reviewed in Bus and Hammond, 2007; Garabrant and Philbert, 2002; Health Canada PACR, 2005, 2007; Munro *et al.*, 1992; U.S. EPA RED, 2005). In a 13-week dietary study in F344 rats, a No Observed Adverse Effect Level (NOAEL) and Low Observed Effects Level (LOEL) of 15 and 100 mg/kg/day, respectively, were established for 2,4-D based on decreases in body weight gain (males), platelet count, and thyroid hormone

levels (T3 and T4), and increases in liver, kidney, and thyroid weights and adrenal cortex hypertrophy (females) (Charles *et al.*, 1996). The overall NOAEL established for 2,4-D from a previous rat 2-generation reproductive toxicity study (Rodwell and Brown, unpublished data; reviewed in Munro *et al.*, 1992) was 5 mg/kg/day. However, this study did not assess several reproductive parameters common to current toxicity testing standards, eg, sperm analysis or endocrine-sensitive endpoints such as anogenital distance (AGD). Although both the acute and chronic adult neurotoxicity of 2,4-D had been examined in rats (Mattsson *et al.*, 1997; acute NOEL = 15 mg/kg, mild locomotor effects; chronic NOEL = 75 mg/kg/day, retinal degeneration), 2,4-D has not been evaluated for DNT in a standardized protocol. Immunotoxicity evaluations in adult mice have not identified 2,4-D-related effects (Blakley, 1986; Blakley and Blakley, 1986; Blakley and Schiefer, 1986; Lee *et al.*, 2001); DIT had not previously been assessed.

Much of the toxicity seen in animal studies is observed at high test doses estimated to be at or above saturation of 2,4-D renal clearance (Bus and Hammond, 2007; U.S. EPA RED, 2005), ie, under conditions of nonlinear toxicokinetics (TK). 2,4-D is readily absorbed from the gastrointestinal tract in both rats and humans and rapidly cleared from the body, unchanged, in urine (Gorzinski *et al.*, 1987; Sauerhoff *et al.*, 1977). The initial elimination phase of 2,4-D in male and female rats is between 30 min and 3.4 h, which accounts for the elimination of essentially all absorbed 2,4-D (Pelletier *et al.*, 1989; van Ravenzwaay *et al.*, 2003). 2,4-D renal clearance is facilitated by an active anion transporter (OAT1) that is saturated at oral gavage doses ≥ 50 mg/kg/day in adult male rats (Gorzinski *et al.*, 1987; Timchalk, 2004; van Ravenzwaay *et al.*, 2003). Renal clearance of 2,4-D is saturated at lower doses in female rats, consistent with androgen-enhanced expression of the OAT1 transporter (Buist *et al.*, 2002; van Ravenzwaay *et al.*, 2003). Nonlinear TK due to renal saturation has also been shown in dietary studies in male F344 rats (Saghir *et al.*, 2006) and in both male and female Crl:CD(SD) rats (Saghir *et al.*, 2013).

Animal toxicity observed under conditions of nonlinear TK is of questionable value for extrapolation to potential human risk if the inflection point for onset of nonlinear TK is substantially above human exposure (Barton *et al.*, 2006; Carmichael *et al.*, 2006; Slikker *et al.*, 2004). In a dietary probe study conducted to characterize TK across life stages and support TK-derived dose selection for the 2,4-D EOGRTS (Saghir *et al.*, 2013), P1 females exhibited nonlinear TK at ≤ 200 ppm (approximately 14 mg/kg/day), whereas P1 male nonlinearity was evident at ≥ 800 ppm (approximately 41 mg/kg/day). Nonlinear TK was seen in postnatal day (PND) 35 pups at 400 ppm (females; lowest dose tested in the probe study) and 800 ppm (males). Importantly, biomonitoring of farm families during active use of 2,4-D has demonstrated that human exposures to 2,4-D are very low and substantially below the inflection points of TK nonlinearity in rats, with geometric mean daily doses determined as ≤ 2.46 μ g/kg/day in applicators and 0.8

and 0.22 μ g/kg/day, respectively, in spouses and in children living on farms where 2,4-D was applied (Alexander *et al.*, 2007; Thomas *et al.*, 2010).

EOGRTS dose levels were selected based on a combined analysis of toxicity and TK data from a 2,4-D probe study and subsequent dose titration study (Saghir *et al.*, 2013). Because of the aforementioned gender differences, different dietary concentrations were selected for adult males and females predicted to achieve high doses equal to or just greater than the inflection point for nonlinear TK. The female high dose was approximately 30 mg/kg/day (600 ppm), whereas the male high dose was approximately 40 mg/kg/day (800 ppm). The low dose of 5 mg/kg/day (100 ppm) was predicted to identify a NOAEL, consistent with the previously identified reproductive toxicity NOAEL (Rodwell and Brown, unpublished data). 2,4-D is adequately transferred through maternal milk to pups (Saghir *et al.*, 2013), so direct gavage dosing of neonates was not needed to assure continued pup treatment during sensitive periods of postnatal development. Dietary concentration adjustments during lactation and the early postweaning period were designed to provide a relatively constant mg/kg body weight/day dose throughout all study phases. F1 male and female offspring were transitioned to adult concentration diets on PND 35, when gender differences in 2,4-D TK may begin to appear (Buist *et al.*, 2002).

The purpose of this EOGRTS was to evaluate the potential effects of 2,4-D on P1 reproduction, and F1 offspring growth and development, including functional assessments of the nervous system, immune system, and endocrine function (including thyroid perturbations) following exposure during critical windows of development. The EOGRTS reserved sufficient animals to conduct a second-generation breeding; however, evaluation of P1 and F1 animals determined this mating to be unnecessary using *a priori* established criteria (see section on Triggers for Second-Generation Mating). The EOGRTS also included an extensive evaluation of systemic toxicity to facilitate data interpretation and establish relative sensitivity of toxicity endpoints.

It is not possible to present the entire 2,4-D EOGRTS data set in this article; therefore, selected data are shown, including data on dosimetry, systemic toxicity, reproductive toxicity, endocrine toxicity (androgen, estrogen, and thyroid pathways), DNT, DIT, and TK. Additional data are presented in the Supplementary Tables. Given the variety of endpoints evaluated, discussion on the significance of these results follows each data set.

MATERIALS AND METHODS

Materials. 2,4-D was obtained from Nufarm Americas, Inc. ($\geq 97.85\%$ pure). The positive control immunotoxicity compounds were cyclophosphamide (CP) monohydrate (Sigma, St Louis, Missouri) and Anti Asialo GM1 (rabbit) (Wako Chemicals USA, Inc., Richmond, Virginia).

Test animals. Animal usage was reviewed and approved by the Institutional Animal Care and Use Committee. Male and female Crl:CD(SD)

rats (Sprague Dawley derived), approximately 9 weeks of age (Charles River Laboratories, Inc., Portage, Michigan), were acclimated to the laboratory (AAALAC International accredited) for 7 days before study initiation. Rats were assigned to treatment groups using a computer program designed to increase the probability of uniform group mean body weights and standard deviations. Animals were uniquely identified using SC implanted transponders (BioMedic Data Systems, Seaford, Delaware).

Housing, water, and diet. Rats were provided Certified Lab Diet no. 5002 (PMI Feeds, St Louis, Missouri) and housed in stainless steel cages or in plastic nesting cages with ground corn cob bedding from late gestation through lactation and the early postweaning period to PND 28. Animals were provided drinking water and diet *ad libitum*.

Dosing. The test material was air milled (particle size: approximately 149 μ m) to minimize clumping and facilitate homogenous diet mixes. A concentrated test material-feed mixture (premix) was prepared and diluted with ground feed to achieve the targeted concentrations. Homogeneity and concentrations of 2,4-D were confirmed by liquid chromatography-mass spectrometry at 4 weeks (end of prebreeding period), 8 weeks (approximately the end of gestation), 12 weeks (at weaning), 15 weeks, and 19 weeks. Test diets ranged from 86.7% to 106.8% of the targeted concentrations with all but 4 of the 55 samples analyzed within $\pm 10\%$ of target. Test diets were used within established stability limits (27 days).

TK and toxicity data were used for dose-level selection (Cooper *et al.*, 2006), with the high dose targeted at or slightly above the threshold for nonlinear TK. Different high-dose levels were selected for males and females based on gender differences in renal clearance (van Ravenzwaay *et al.*, 2003). The male high dose (800 ppm, approximately 45 mg/kg/day) was slightly higher than the inflection point for nonlinear TK in male pups from PND 35 to adulthood, whereas the female high dose (600 ppm, approximately 40 mg/kg/day) was clearly higher than the TK inflection point in female pups and adults (Saghir *et al.*, 2013). Dietary concentrations were adjusted during lactation and pup postweaning periods using recent body weight/feed consumption data or historical control data with consideration of the large feed consumption increases during specific life stages (Table 1; Carney *et al.*, 2004; Saghir *et al.*, 2013). These adjustments provided a relatively constant mg/kg body weight/day dose across life stages. On PND 35, F1 offspring were returned to adult dietary concentrations of 2,4-D until termination.

Study design. The study design is illustrated in Figure 1. Supplementary Table S1 shows the endpoints evaluated in each generation of rats. Beginning 10 weeks of age, rats (27/sex/dose) were fed diets containing 0, 100, 300, and 600 (females) or 0, 100, 300, and 800 (males) ppm 2,4-D, supplying approximately 0, 7, 21, or 40 mg/kg/day 2,4-D for adult females and 0, 6, 17, or 45 mg/kg/day 2,4-D for adult males for approximately 4 weeks prior to breeding and continuing through breeding (up to 2 weeks), gestation (3 weeks), and lactation (3 weeks). Exposure of P1 males continued for an additional 7 weeks after the initiation of the mating phase to ensure coverage of a full spermatogenic cycle. P1 females were exposed until lactation day (LD) 22 (the end of the lactation

period). Selected F1 offspring were maintained on the test diet until PND 60 (Set 1b, 10/sex/dose), approximately PND 70 (Sets 1a, 10/sex/dose, and 2a, 10/sex/dose), approximately PND 90 (Set 2b, 10/sex/dose), and approximately PND 139 (Set 3, 20/sex/dose).

P1 males and females were evaluated for systemic toxicity and for functional and structural evaluations of the reproductive systems. Results from *a priori* defined reproductive and toxicity endpoints from the P1 and F1 generations were assessed to determine whether production of a second generation was triggered (see Results section). In addition, a satellite group of P1 females (12 per dose) was exposed during prebreeding, breeding, and gestation days (GD) 0–17, when they were euthanized to assess gross pathology, clinical pathology, thyroid hormones, TK, and selected reproductive parameters (ie, female mating/fertility/conception indices; time to mating; numbers of corpora lutea, implantations, and resorptions; viable fetuses/litter; and percent pre- and postimplantation losses) during gestation.

F1 offspring were evaluated for effects on systemic toxicity (Set 1a), nervous system (Set 1b), immune system (Set 2a and 2b), reproductive system, and thyroid function (Set 3). 2,4-D TK was assessed in the Set 3 F1 offspring on PND 63 and 84. Unselected weanlings were sacrificed on PND 22 and evaluated for systemic toxicity and neuropathology. The various study designs (parental group, satellite group, Sets 1–3 and unselected weanlings) are illustrated in Figure 1 and Supplementary Figures S1–S5.

Clinical observations. All animals were observed twice daily for alterations in behavior or demeanor and were given weekly handheld clinical observations. In addition, detailed clinical examinations, using categorical and ranked (ie, scored) observations, were conducted on the P1 animals preexposure and once during the last week of the premating period.

Body weights and feed consumption. P1 male body weights were collected weekly throughout the study. P1 females were weighed weekly prior to breeding, on GD 0, 7, 14, and 20, and LD 1, 4, 7, 14, and 21. The body weight of each pup was recorded on PND 1, 4, 7, 14, and 21 and weekly after weaning until termination. Mated females from the satellite group were weighed weekly during the preexposure and premating periods, and then on GD 0, 7, 14, and 17. Body weight gains were calculated for corresponding intervals. The Set 1b F1 offspring also were weighed on the day of the functional observational battery (FOB) and acoustic startle response (ASR) measurement. Generally, feed consumption measurements were determined for the same intervals over which body weights were collected. Test material intake (TMI, expressed as mg/kg/day) was calculated at designated intervals.

Estrous cycle (P1 females and F1 Set 3 females). Estrous cycle length and normality were evaluated by daily vaginal lavage in P1 females for 2 weeks prior to breeding and in Set 3 F1 females for 4 weeks from PND 40–68.

Breeding procedure (P1 animals). In the main study, each P1 female was placed with a single P1 male from the same dose level (1:1 mating), whereas in the satellite group, each female was placed with an untreated male that had access to 2,4-D-containing diet only during the mating period. Animals

TABLE 1
2,4-D Dietary Adjustments During Lactation and Postweaning

Exposure Period	TMI Increase ^{a,b}	Unadjusted Concentration (ppm)	Adjustment Factor	Adjusted Concentration (ppm)
LD 7–14	3.1×	100, 300, 600	2	50, 150, 300
LD 14–21	3.8×	100, 300, 600	3	33, 100, 200
PND 21–28 ^c	2.4×	100, 300, 600	2	50, 150, 300
PND 28–35 ^d	1.9×	F: 100, 300, 600 M: 100, 300, 800	2	F: 50, 150, 300 M: 50, 150, 400

Note. ^aTMI, test material intake is based on feed concentrations, body weights, and feed consumption.

^bRelative TMI compared to nonpregnant adult females.

^cTMI based on feed consumption data derived from male Crl:CD(SD) rats at PND 23–28 from Marty *et al.* (2003).

^dAfter PND 35, pups returned to unadjusted dietary concentrations.

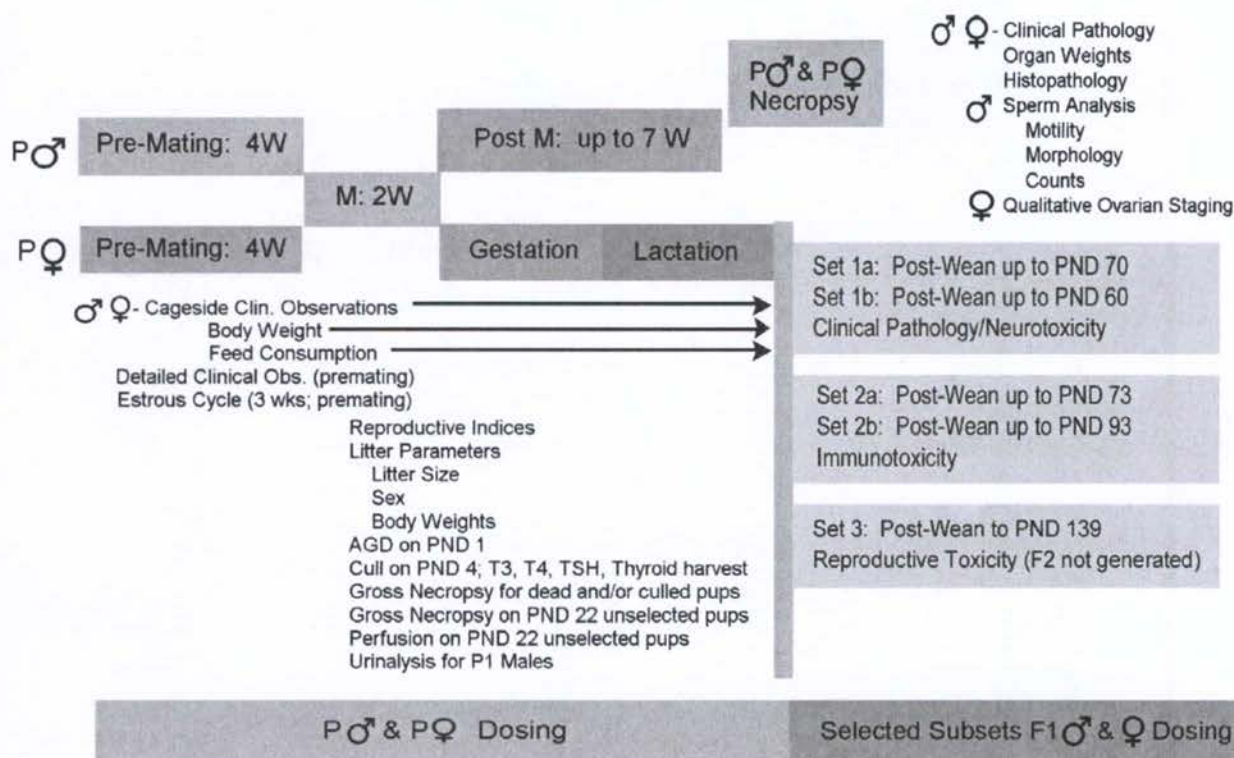


FIG. 1. Study design for the 2,4-D EOGRTS. Abbreviations: W, weeks; M, mating.

were paired until mating occurred or 2 weeks elapsed. The day on which sperm were detected or a vaginal copulatory plug was observed *in situ* was considered GD 0.

Litter data and culling (F1 offspring). For each litter, data included the date of parturition; the number of live and dead pups on LD 0, 1, 4, 7, 14, and 21; and the sex and body weight of each pup on LD 1, 4 (before and after culling), 7, 14, and 21. To minimize variation in pup growth, litters with a total number of pups exceeding 10 were culled to 5 males and 5 females, where possible, on LD 4. Culled pups were randomly selected, examined grossly, and euthanized.

Weaning and pup assignment (F1 offspring). All litters were weaned on LD 21. Three male and 3 female F1 pups/litter were selected randomly, and 1 male and 1 female each was assigned to Set 1, 2, or 3. For Set 1, 1 male or 1 female per litter was assigned to Set 1b, so that the DNT group would include representatives of each litter and control for litter effect prior to statistical analyses ($n = 10$ males + 10 females/dose level). The corresponding male or female from each litter was assigned to F1 Set 1a systemic toxicity. For Set 2, 1 male and 1 female per litter were assigned to Set 2a for the initial evaluation of immunotoxicity with the SRBC AFC assay and then were assigned to Set 2b for the natural killer cell assay ($n = 10$ males + 10 females/dose level with 10 litters represented in each assay). The natural killer cell assay was identified *a priori* for Set 2b animals if the SRBC assay was judged to be negative, which was consistent with results obtained by Cooper *et al.* (2006). For Set 3, 1 male and 1 female per litter were randomly selected for reproductive toxicity assessment ($n \geq 20$ males and 20 females/dose level).

From the unselected F1 weanlings, 1 PND 22 pup/sex/litter ($n = 12$ /sex/dose) was perfused for neuropathology evaluations. Additionally, 1 PND 22 pup/sex/litter ($n = 10$ /sex/dose) was randomly selected for necropsy, with the collection of blood for thyroid hormone analyses, organ weights, and preservation of tissues for subsequent histopathological examination. Remaining unselected F1 weanlings were examined grossly on PND 22.

Endocrine endpoints (F1 offspring prior to maturity). AGD (absolute and relative to the cube root of body weight) was measured in all F1 pups on PND 1 (Gallavan *et al.*, 1999). All F1 offspring were evaluated for the presence of nipple/areolae on PND 12 as described by McIntyre *et al.* (2001). All F1 animals retained postweaning (3/sex/litter assigned to Sets 1–3) were examined daily for vaginal opening beginning PND 26 (Cooper *et al.*, 1989) or for balanopreputal separation beginning PND 35 (Korenbroet *et al.*, 1977), continuing until these markers for puberty were attained. Age and body weights of the animals were recorded on the day puberty was achieved. For each dose group, the mean age and body weight at puberty onset was calculated from the litter means. For AGD, nipple retention, and puberty onset measurements, observers were blind to treatment group when evaluating the F1 offspring.

Thyroid assessments. Rats were anesthetized with isoflurane and examined grossly, and blood was collected from PND 4 culled pups (10 litters/dose; samples pooled by sex and litter with 1 male and 1 female sample/litter), unselected PND 22 weanlings (nonperfused), GD 17 satellite dams, and Set 1a F1 offspring (10/sex/dose, approximately PND 65). Animals were not fasted prior to these blood collections to avoid the potential effects of fasting on thyroid hormone levels (Boelen *et al.*, 2008; Döhler *et al.*, 1979). Serum was analyzed at AniLytics (Gaithersburg, Maryland) for triiodothyronine (T3), thyroxine (T4), and thyroid-stimulating hormone (TSH) using radioimmunoassays with standards from the National Hormone and Pituitary Program (UCLA, California). At necropsy, thyroid/parathyroid glands were harvested from the same animals used for thyroid hormone analyses.

TK (GD 17 satellite dams, Set 3 F1). Blood samples (5 per dose group) collected from nonfasted GD 17 dams at necropsy at approximately the same time each day were used for single-time point TK analyses to determine plasma levels of 2,4-D. For Set 3 adults, 3 blood samples (6 AM, 9 AM, and 5 PM per Saghir *et al.*, 2006) were collected from randomly selected nonfasted animals (5 rats/sex/dose group) on PND 63 and PND 84 for the determination

of 2,4-D plasma area under the curves (AUCs). Stability of 2,4-D in plasma was established previously (Saghir *et al.*, 2013).

Clinical pathology (P1 GD 17 satellite dams, Set 1a F1 offspring). Animals (10/sex/dose) were anesthetized with O₂/CO₂ at scheduled necropsy, and blood samples were collected from the orbital sinus from fasted P1 animals (males after approximately 11 weeks of treatment, females on LD 22). Set 1a F1 animals (approximately PND 70), and nonfasted P1 GD 17 satellite dams. Hematology parameters, coagulation, clinical pathology analytes, urinalysis, and thyroid hormones were evaluated using instrumentation/methods listed in Supplementary Table S2.

DNT assessments (Set 1b F1 offspring): Set 1b F1 offspring (10 males and 10 females representing 20 litters) received a FOB between PND 54 and 56 at the same approximate time each test day. The FOB, which was conducted according to previously described procedures (Mattsson *et al.*, 1986, 1997), included cage-side, handheld, and open-field observations with ranked and categorical observations, as well as sensory evaluations, which included tests for nociception (responsiveness to tail pinch) and startle response (responsiveness to sharp noise), and measurements of body weight, rectal temperature, forelimb and hind limb grip performance, landing foot splay, and motor activity (MA). The FOB was conducted by an observer who was blind to the treatment status of the animal. The same observer was used for all rats. Each MA session consisted of eight 8-min epochs, totaling 64 min of testing per animal per test session as determined in a validation study (Marty and Andrus, unpublished data). ASR habituation was evaluated between PND 57 and 59 using a commercially available ASR system (Med Associated, Inc., St Albans, Vermont) under the following conditions (Andrus, unpublished data): approximately 65 dB[A] background noise, 50 startle trials (approximately 120 dB[A] white noise burst, 50-ms duration, 2-ms rise/fall time with 10–20 s variable intertrial intervals) with mean peak response amplitude calculated for each 10-trial blocks. Body weights were collected on the day of ASR assessment. Equipment used for rectal temperature, grip performance, MA, and ASR were calibrated before testing.

DIT assessments (Set 2 F1 offspring). F1 offspring in Set 2a (10/sex/dose using 1/sex/litter) were evaluated for immune function using the sheep red blood cell (SRBC) antibody plaque forming cell (AFC) assay (Ladics *et al.*, 2000). On PND 66–70, F1 offspring were immunized by IV injection with SRBC. AFC response was evaluated 4 days postimmunization as described by Carney *et al.* (2004). Five animals per sex from similarly aged satellite animals served as a concurrent positive control group (immunized with 4 daily doses of 20 mg/kg/day CP via IP injection as described by Loveless *et al.* (2007)).

Set 2b F1 offspring (10/sex/dose using 1/sex/litter) were subjected to a natural killer cell (NK) assay (Marcusson-Stahl and Cederbrant, 2003). Spleens were removed from Set 2b animals (PND 87–93), and a single-cell suspension of splenocytes was prepared. Splenocytes (effectors) were incubated with previously plated target cells (mouse lymphoma cell line YAC-1 from ATCC [Manassas, Virginia] labeled with carboxyfluorescein succinimidyl ester [CFSE]) at effector:target ratios that range from 6.25:1 to 200:1. At the end of incubation (4 h, 37°C), propidium iodide (PI) was added to identify dead cells, and samples were analyzed by flow cytometry (Beckman Coulter Epics XL). A separate group of age-matched rats received an IV injection of rabbit anti-mouse/rat Asialo GM1 polyclonal antibody 24 h prior to necropsy (positive control group). Comparative assessments of the flow cytometry and chromium release approaches indicate a similar sensitivity in their ability to detect NK cell activity and inhibition (Cederbrant *et al.*, 2003; Kim *et al.*, 2007; Langhans *et al.*, 2005). Enumeration of lymphocyte subpopulations was not conducted in the 2,4-D EOGRTS, although this endpoint was subsequently included in the OECD 443 test guideline.

Pathology. For most adult P1 and all F1 females, vaginal lavage samples were collected to confirm estrous stage at the time of termination (lavage samples were inadvertently not collected on 6 to 7 P1 females/dose during the first 3 days of P1 necropsy). The uteri of all P1 females were stained with 10% sodium sulfide stain to count implantation sites. Animals were anesthetized with O₂/CO₂ and weighed, and when appropriate, blood samples were collected

(TK, thyroid hormone, clinical chemistry/hematology). Animals were euthanized by decapitation. A complete necropsy was conducted on all animals by a veterinary pathologist assisted by a team of trained individuals with the exception of Set 2 animals, which had a limited necropsy. Where designated, organs were collected and weighed, and ratios of organ weight-to-terminal body weight were calculated. The organs weighed, tissues preservation methods, stains, and tissues examined for each group are listed in Supplementary Table S1. Preserved tissues from the high-dose and control animals were prepared by standard techniques, sectioned approximately 5–6 µm thick, and examined by a veterinary pathologist using a light microscope. Histopathologic findings were subjectively graded as appropriate to assess the potential effects of exposure with regard to the contribution of a specific lesion to the health status of an animal; tissues from the lower dose groups were only examined if exposure-related effects were seen in the high-dose group, gross lesions were present, or the animals failed to mate and/or produce a litter (P1 males and females only).

Sperm parameters (motility, morphology, and testicular sperm/epididymal sperm counts) were evaluated using the HTM Integrated Visual Optical System (IVOS; Hamilton-Thorne Research, Beverly, Massachusetts) as described by Carney *et al.* (2004). Sperm motility was examined in P1 and Set 3 F1 males from all dose groups. Sperm morphology and count data were generated for control and high-dose P1 and Set 3 males, as well as for any males that failed to mate successfully during the mating period. Because of the absence of exposure-related effects, sperm from the lower dose levels were not evaluated.

Perfusions for unselected PND 22 weanlings and Set 1b F1 offspring. Weanlings selected for perfusion and Set 1b F1 adults (PND 60) were given an IP injection of 0.2 ml heparin (10 000 USP/ml) per 100 g body weight, anesthetized with O₂/isoflurane, cannulated via the left ventricle of the heart, and perfused by gravity pressure (height was 66.6 inches above the animal to produce a pressure of 120 mm Hg) with 0.05M phosphate buffer containing sodium nitrite followed by a PBS of 1.5% glutaraldehyde-4% formaldehyde (c. 540 mOs). Tissues were examined for gross pathologic alterations by a veterinary pathologist. Brain weight and gross linear measurements were recorded on all dose groups. Tissues included in the neuropathology assessment are described in Supplementary Table S1.

For morphometric purposes, 2 transverse tissue blocks were cut through the cerebrum and midbrain (blocks 3 and 4), and a longitudinal (anterior to posterior) cut was made midway through the cerebellum after it was removed from the midbrain (block 10). Blocks 3 and 4 contained the cerebrum (frontal and parietal lobes), thalamus/hypothalamus, and midbrain. These tissues were processed by standard histologic procedures. Sections that met the criteria for appropriate microscopic landmarks were used to measure specific microscopic structures. Images of each brain section were digitally captured using an Olympus DP70 or a SPOT RT camera mounted on a Leitz dissecting microscope (blocks 3 and 4) or a Leitz light microscope (block 10) with image capture software (Diagnostic Instruments, Inc., Sterling Heights, Michigan). Simple morphometric measurements were obtained using the Image-Pro Plus image analysis software (Media Cybernetics, Silver Spring, Maryland). Measurements were excluded as necessary due to microscopic artifacts, such as missing portion(s) of tissue and tears in tissue. Additional sections from blocks 3, 4, and 10 from the brain were stained with Luxol Fast Blue to assess myelin. Evaluations were only conducted in the high-dose and control animals as there were no exposure-related findings.

Animals examined at unscheduled necropsy. A complete set of tissues (listed in Supplementary Table S1) were examined histopathologically from one 300 ppm Set 3 male rat found dead on PND 89. One 300 ppm Set 3 F1 female died on PND 24 prior to scheduled necropsy. Tissues from this female were inadvertently not examined histopathologically. Deaths of these 2 animals were attributed to accidental injuries.

Statistics and calculations. Statistical analyses used for this study are outlined in Supplementary Table S3. Where appropriate, the litter was the experimental unit of analysis. The type I error rate (alpha) was set to .05 for all primary planned analyses. Reproductive indices and SRBC responses were

calculated using standard equations (Supplementary Table S4). Nonpregnant females, females with resorptions only, or females found to be pregnant only after staining of their uteri were excluded from subsequent gestation and/or lactation parameters. Many analyses were performed separately by sex due to the recognized gender differences in 2,4-D response and differences in the dose levels used for males and females.

For TK data, descriptive statistical (ie, mean \pm SD) analyses were performed using Microsoft Excel (Microsoft Corporation, Redmond, Washington) spreadsheets in the full precision mode (15 digits of accuracy). Individual values obtained were analyzed for AUC₀₋₂₄ by trapezoidal method using PK Functions for Microsoft Excel (PK Functions for Microsoft Excel, Joel I. Usansky, Ph.D., Atul Desai, M.S., and Diane Tang-Liu, Ph.D., Department of Pharmacokinetics and Drug Metabolism, Allergan, Irvine, California; <http://www.boomer.org/pkin/soft.html>), a method previously used by other investigators (Beringer *et al.*, 2005; Pizarro *et al.*, 2004). Linearity was determined graphically by regression analysis (as per Roberts and Renwick, 1989; Sweatman and Renwick, 1980) and by comparing ratios of plasma 2,4-D and AUC_{24h}.

RESULTS AND DISCUSSION

Dosimetry

Using body weight and feed consumption data, TMI for the various study phases was determined (Table 2). Targeted dose levels for this study were 5, 15, and 30 mg/kg/day (females) and 5, 15, and 40 mg/kg/day (males). The TMI values for the P1 males were close to the predicted values. However, in the P1 females, TMI exceeded the targeted dose levels, particularly during the early lactation period (LD 1–7), when TMI was double the targeted dose levels. To limit overexposure of the offspring during critical windows of development in accordance with guidance by Cooper *et al.* (2006), dietary concentrations of 2,4-D were adjusted during subsequent lactation phases (LD 7–21) to account for the large and rapid increase in feed consumption (3- to 4-fold) typical for rats in mid- to late lactation (Carney *et al.*, 2004). These dietary adjustments were intended to maintain more consistent doses of 2,4-D (mg/kg/day basis) across study phases. Despite adjustments in the dietary concentrations of 2,4-D, doses for main study females were still greater than target levels during the second week of lactation (LD 7–14). TMI is not reported for the last week of lactation (LD 14–21) as pups also consume diet during this period (Hanley and Watanabe, 1985; Yoon and Barton, 2008), making dam TMI calculations inaccurate.

This study also included dietary concentration adjustments for the F1 offspring until PND 35; thereafter, F1 offspring received significantly larger doses than targeted. The highest TMIs were from PND 35–42, just after dietary adjustment ended (Supplementary Table S5). The longer the F1 animals were maintained on test diet beyond PND 70, the closer TMI values approached the targeted dose levels. Thus, animals in Set 1b, which were euthanized on PND 56 (earliest time point), had the highest average 2,4-D doses, whereas Set 3, euthanized on PND 139, had 2,4-D doses closer to targeted levels. In Sets 1a, 1b, and 2a, all of which were euthanized by PND 70, doses were nearly twice the targeted dose levels of 5, 15, and 30 or 40 mg/kg/day.

TABLE 2
TMI (mg/kg/day) for P1 Males and Females and F1 Offspring

Nominal Dose (ppm)	P1 Males and Females						F1 Offspring ^a									
	Males (TD 1–78)	Females (TD 1–29)	Main Study Females (GD 0–20)	Satellite Females (GD 0–17)	Lactating Females (LD 1–7)	Lactating Females (LD 7–14) ^{b,c}	Set 1a (PND 28–69)		Set 1b (PND 28–56)		Set 2a (PND 28–70)		Set 2b (PND 28–84)		Set 3 (PND 28–133)	
							M	F	M	F	M	F	M	F	M	F
100	5.51	6.97	7.37	7.63	10.7	8.4	9.24	9.56	9.88	10.1	9.15	9.66	8.67	9.05	6.83	7.59
300	16.6	20.6	21.9	22.9	32.4	25.5	28.4	28.8	29.5	30.0	28.4	28.7	25.8	26.7	20.9	23.3
600	—	40.2	43.7	45.3	64.5	51.5	—	57.9	—	59.2	—	58.4	—	55.3	—	46.7
800	45.3	—	—	—	—	—	76.6	—	81.7	—	75.3	—	71.8	—	55.6	—

Note. —, not applicable; TD = Test day.

^aF1 males and females did not receive the adult dietary concentration until PND 35; TMI calculations were corrected for adjusted diets given on PND 28–35.

^bDietary concentrations were decreased by one-half during the LD 7–14 interval to account for increased feed consumption by the dams; TMI calculations were corrected for the adjusted diets.

^cTMI was not calculated for LD 14–21 as pups also are consuming diet during this interval.

Whereas lactating dams received a greater dose/kg body weight than was administered during the premating period, 2,4-D-treated dams and pups would have received much higher doses during these critical periods of development if these adjustments were not made (Saghir *et al.*, 2013). Dietary concentration adjustments during lactation were justified as nursing humans do not experience the large increase in feed consumption seen in rats. In humans, the metabolic demands of breastfeeding require an increase in maternal metabolism of approximately 20% or less (Frigerio *et al.*, 1991; Prentice and Whitehead, 1987). This contrasts markedly with laboratory rodents, where nutrient intake is critical for lactation success because a large portion of the metabolic output is directed to milk synthesis (Grigor *et al.*, 1987; Smart *et al.*, 1987).

Dietary adjustments for the F1 offspring continued until PND 35; however, the F1 animals consumed greater amounts of feed per kilogram body weight for a much longer period beyond PND 35 and did not reach adult feed consumption levels per kilogram body weight until approximately 10 weeks of age, the age at which dosing began in the P1 animals. It is difficult to maintain consistent dosimetry during this period because the animals are growing rapidly with rapid changes in body weight:feed consumption ratios. Although children also consume greater amounts of food per kilogram body weight, dietary concentration adjustments during this period are warranted for 2 reasons. First, when toxicity was seen, it was difficult to discern if this toxicity was related to high peak exposures during the juvenile period or whether exposure to the targeted concentrations produced the effects. With dietary adjustments during the juvenile period, an appropriate NOAEL can be identified across life stages; the increased food consumption in children must be (and is) considered in the exposure component

of risk assessment to ensure that this toxicity threshold is not exceeded. Second, the increased dose received by these immature animals exceeded the threshold for nonlinear kinetics for several weeks at both the mid- and high doses, a condition that would not occur in humans with 2,4-D due to large margins of exposure (see TK section). Prolonged exposure to doses in the nonlinear TK range likely contributed to the more pronounced kidney toxicity seen in F1 offspring compared with P1 animals (see Systemic Toxicity section).

The longer the F1 animals were maintained on test diet beyond PND 70, the closer TMI values approached the targeted dose levels. These results indicate that it would have been appropriate to extend the period of dietary dose adjustments beyond PND 35 to lessen the extent of overexposure during this developmental period. However, only a finite number of dietary concentrations can be managed logistically during an EOGRTS study. Additionally, for practical reasons, the range-finding study was terminated at PND 35; thus, there were no TK data available to predict the overexposure without dietary adjustment after PND 35.

TK

Dietary 2,4-D TK was extensively evaluated in the range-finding study (Saghir *et al.*, 2013). Additional TK was included in the current study to confirm that gestational TK did not differ from other life stages and to extend 2,4-D TK data in the F1 offspring to later life stages. Plasma 2,4-D levels were measured in GD 17 satellite females and in F1 Set 3 males and females on PND 63 and 84 (Table 3). The TMI (mg/kg/day) in Table 3 shows the dietary intake (dose in mg/kg/day) of 2,4-D during the week prior to blood collection; therefore, these values differ somewhat from Table 2, which shows cumulative dose (mg/kg/day) for these groups.

TABLE 3
TK in Satellite Females (GD 17) and F1 Male and Female Set 3 Offspring (PND 63 and 84)

Test Group	Dose (ppm)	Dose (mg/kg/day)	Fold Increase	2,4-D Concentration or AUC _{24h} ^{a,b}	Fold Increase
GD 17 dams	100	6.97	1×	1.21	1×
	300	21.89	3×	4.71	4×
	600	42.57	6×	39.9	33×
PND 63 males	100	7.81	1×	18.60	1×
	300	22.0	3×	76.51	4×
	800	59.6	8×	302.24	16×
PND 63 females	100	8.39	1×	38.66	1×
	300	26.1	3×	236.15	6×
	600	49.0	6×	1380.03	36×
PND 84 males	100	5.89	1×	15.73	1×
	300	18.4	3×	62.08	4×
	800	49.4	8×	234.89	15×
PND 84 females	100	7.19	1×	26.78	1×
	300	21.3	3×	123.59	5×
	600	46.6	6×	500.54	19×

Note. Bolded cells represent doses that exhibited nonlinear TK.

^aGD 17 dam values are single point measurements (μg/g), whereas F1 offspring measurements were AUC_{24h} (μg h/ml).

^bMean values for 2,4-D concentration (GD 17 dams) or AUC_{24h} (F1 offspring); n = 4 per dose.

Based on proportional TMI, the increase in the systemic dose (AUC_{24h}) during gestation at the middle-dose level (300 ppm; 22 mg/kg/day) was slightly above expected levels (4 times vs predicted 3 times). This small increase above the predicted systemic dose means that the 300 ppm dose level was at or slightly above linear TK during the GD 17 period. The observation of possible nonlinear TK at 300 ppm (22 mg/kg/day) in GD 17 females also was consistent with TK findings in nonpregnant females in the range-finding study in which onset of nonlinear TK likely occurred at a dietary dose \leq 200 ppm (approximately 14 mg/kg/day) (Saghir *et al.*, 2013). The systemic dose of 2,4-D became highly nonlinear at the highest concentration of 600 ppm (43 mg/kg/day) in GD 17 females. If linear TK is assumed, the dose-dependent increase in plasma concentrations expected in the high-dose 600 ppm group was approximately 7.26 μ g/g (6-fold increase relative to the 100 ppm systemic dose); however, the actual systemic dose was 39.9 μ g/g, a 33-fold increase. Thus, the TK data indicated that the high-dose treatment in GD 17 rats (43 mg/kg/day) was substantially above the inflection point identified for onset of nonlinear TK (approximately 22 mg/kg/day).

In PND 63 F1 males, the systemic AUC dose of 2,4-D was increased 16-fold at 800 ppm (8-fold expected for linearity) and 4-fold at 300 ppm (3-fold expected for linearity), making the 300 ppm dose at or slightly above the threshold for nonlinear TK. Similarly, PND 84 males had a 15-fold increase in AUC_{24h} at the high dose (8-fold expected) and a 4-fold difference at the mid dose (3-fold expected), indicating that the threshold for nonlinear TK in F1 males was approximately 300 ppm. Female PND 63 F1 animals exhibited 6- and 36-fold systemic 2,4-D increases for the respective 300 and 600 ppm doses (3- and 6-fold expected for linearity), indicating that nonlinear TK behavior was present even at the middose level. In PND 84 F1 females, systemic doses were increased 5- and 19-fold, respectively, at the mid- and high-dose levels (3- and 6-fold expected for linearity), again indicating that these dose levels were slightly above the inflection point for nonlinear TK in F1 females.

Overall, the TK evaluations confirmed that the high-dose levels selected for this study (600 ppm females; 800 ppm males) achieved and likely exceeded a primary study objective of producing systemic plasma concentrations at the high doses that were at or slightly above the threshold exhibiting nonlinear TK behavior. One advantage of this study design is that it promotes the use of TK data to set a kinetically derived maximum dose (KMD) compared with use of the traditional maximum tolerated dose (MTD) approach to define study dose selection (Saghir *et al.*, 2012, 2013). An MTD-based dose selection strategy based primarily on body weight and/or other evidence of toxicity would likely have resulted in selection of mid and high study doses higher than the threshold dose for onset of nonlinear TK in rats (Saghir *et al.*, 2013). Thus, an MTD-based dose selection strategy would have almost certainly resulted in at least 2 of the 3 study doses producing internal systemic

dose conditions that were not relevant to assessing human health risks, and would have compromised generation of dose-response data that are most informative of actual human health risk.

As described by Cooper *et al.* (2006) and the subsequent OECD test guideline, a dose at or slightly above the threshold for nonlinear TK is a suitable high-dose level, provided that there is a reasonable margin of exposure (MOE) to human exposures. Assuming that the inflection for onset of nonlinear TK is approximately 20 mg/kg/day in females (Table 3; Saghir *et al.*, 2013), and given that high-quality biomonitoring of farm workers and families indicates geometric mean 2,4-D doses of \leq 2.46 μ g/kg/day for applicators and 0.8 μ g/kg/day for nonapplicator spouses (Alexander *et al.*, 2007; Thomas *et al.*, 2010), the respective TK inflection point dose divided by human exposure levels provides MOEs ranging from 8130 to 25 000. In addition, recent "biomonitoring equivalent" analyses (estimates of urine and/or blood concentrations if humans were exposed to the EPA regulatory Reference Dose (RfD) for 2,4-D) have confirmed that actual human exposures are well below the 2,4-D RfD health standard that is conservatively set to be protective of human health (and is derived from the chronic study NOEL of 5 mg/kg/day, which is equivalent to the low dose in this study) (Aylward *et al.*, 2010; Hays *et al.*, 2012). Thus, human exposures to 2,4-D are very disparate from the inflection point of the onset of TK nonlinearity, strongly indicating that results at doses exhibiting nonlinear TK are irrelevant for risk assessment purposes.

2,4-D is eliminated from the body via the renal OAT1 transporter, which is not fully developed in immature rats (Buist *et al.*, 2002). Expression of OAT1 is immature up to approximately PND 35 in both male and female rats (Buist *et al.*, 2002) and thus can result in higher systemic doses in immature animals relative to adults treated with equivalent 2,4-D doses (Saghir *et al.*, 2013). However, the overall NOAEL of 5 mg/kg/day observed for toxicity in this study (see Systemic Toxicity), which is inclusive of young animals, indicates that the KMD approach for dose selection used in this study has provided data with improved relevance for risk assessment while assuring health protection. This is particularly so, given that biomonitoring of farm family children identified a geometric mean dose to children of 0.22 μ g/kg/day (Alexander *et al.*, 2007), resulting in an MOE of 22,727 to the 5 mg/kg/day low dose in this study.

Systemic Toxicity

Body weights. In accordance with the EOGRTS study design, systemic toxicity was assessed across life stages. In the parental generation, high-dose females showed decreased body weight during lactation prior to dietary adjustments (LD 1–7), as did high-dose F1 pups (data not shown). Overall, the F1 pup body weights recovered to control values by approximately PND 42 (females) and 56 (males), although 1 subset of F1 males (Set 1a) had a 10% decrease in terminal body weights at necropsy on PND 70. These findings were considered exposure

related, although these effects occurred at doses in the nonlinear TK, which are not relevant for human risk assessment.

P1 and F1 adult renal toxicity. This study confirmed that the kidneys are a target organ for 2,4-D toxicity; kidneys were affected at similar dose levels/exposure durations to previous 2,4-D studies (Gorzinski *et al.*, 1987; Saghir *et al.*, 2013; Schulze, unpublished data). Absolute and/or relative kidney weights were increased in P1 males in the 800 ppm group. Kidney weights were significantly increased in F1 Set 1a females at 300 ppm only at PND 70, but were not elevated with extended dosing out to PND 139. At 600 ppm, kidney weights were increased in Set 1a and Set 3 females (Table 4) but were not affected in the F1 males even at the high-dose level. In the P1 generation, the characteristic renal lesion, very slight to slight degeneration of the proximal convoluted tubules in the outer zone of the medulla, was seen in high-dose P1 males only. Renal lesions extended to middose F1 adult males at PND 70 and PND 139. Kidney lesions also were present in high-dose F1 adult females. In both generations, renal lesions were slightly more severe in males than in females. No renal lesions were seen in PND 22 F1 pups.

The very minor renal effects seen at 300 ppm were judged not to be treatment related. Although there was some elevation in the incidence of "very slight" kidney toxicity in middose PND 70 males relative to controls (6/10 vs 1/10), 10/10 PND 139 control rats also had an increased incidence of "very slight" kidney toxicity. Thus, characterization of the middose, "very slight" response as treatment related, particularly given the minimal evidence of an increase in severity (3/10 "slight"

at PND 139), was judged questionable. The absence of effect in the middose PND 139 rats points to a female LOEL of 600 ppm and a male LOEL of 800 ppm for kidney toxicity.

The incidence of the kidney histopathological changes in the F1 adults compared with P1 adults was likely related to higher 2,4-D doses in F1 offspring (Tables 2 and 4) and was associated with nonlinear TK at the high-dose level (Table 3). For example, even with dose adjustments implemented between PND 21 and 35, the peak doses of 2,4-D between weaning and termination were 78 and 76 mg/kg/day in the high-dose Set 1a and Set 3 females, respectively, compared with 42 mg/kg/day in prenatally P1 females. Thus, the differences in incidence and degree of these histopathological changes across generations are explained by the 2,4-D dosimetry and TK, and are not attributable to an intrinsic sensitivity of young animals versus adults. These kidney lesions were slight to very slight in severity and would not be expected to alter renal function, which was confirmed by the lack of exposure-related effects on clinical pathology or urinalysis (Supplementary Table S6).

Although the same renal lesions were seen in both males and females, higher dose levels and/or longer exposures were needed to produce these effects in females compared with 2,4-D-exposed males. The reason for the gender-related sensitivity to kidney histopathology is likely related to 2,4-D TK. The OAT1 transporter responsible for excretion of 2,4-D in urine is located on the basolateral membrane of the proximal tubules of the kidney (Hasegawa *et al.*, 2002, 2003; Timchak, 2004; Tojo *et al.*, 1999). Because the expression of OAT1 is higher in males than in females (Buist *et al.*, 2002), the transport of 2,4-D into the proximal tubule from plasma will saturate

TABLE 4
Kidney Alterations in P1 and F1 Offspring Treated With 2,4-D

Sex	Males				Females			
	0	100	300	800	0	100	300	600
P1 (~PND 147)								
Absolute kidney weights (g) ^a	3.513	3.621	3.629	3.965^b	2.328	2.275	2.417	2.395
Relative kidney weights (g/100 g body weight)	0.662	0.686	0.685	0.734^b	0.789	0.766	0.807	0.809
Histopathology: Very slight degeneration	9	4	6	3	0	0	0	0
Histopathology: Slight degeneration	1	0	1	8	0	0	0	0
Set 1a (PND 70)								
Absolute kidney weights ^a	3.108	3.004	3.208	3.050	1.749	1.812	1.946	1.905
Relative kidney weights	0.819	0.803	0.850	0.899	0.765	0.793	0.846 ^b	0.846^b
Histopathology: Very slight degeneration	1	1	6	4	1	0	1	5
Histopathology: Slight degeneration	0	0	0	5	0	0	0	0
Set 3 (PND 139)								
Absolute kidney weights ^a	4.005	4.171	4.167	4.094	2.080	2.146	2.199	2.241^b
Relative kidney weights	0.693	0.678	0.717	0.712	0.698	0.720	0.726	0.751
Histopathology: Very slight degeneration	10	10	8	4	0	0	1	7
Histopathology: Slight degeneration	0	0	3	6	0	0	0	0

Note. Bolded numbers indicate effects interpreted to be exposure related.

^an = 24–27/sex/dose level for P1 kidney weights; n = 10/sex/dose level for Set 1a kidney weights; n = 24–27/sex/dose level for Set 3 kidney weights.

^bStatistically different from control mean by Dunnett's test ($\alpha = .05$).

^cExamined 10–11/sex/dose level; data indicate number of affected animals per treatment group for each observation.

at a higher systemic concentrations in males than in females and thus allow for a greater delivered dose to proximal tubule cells in males. Put the opposite way, saturation of OAT1 at lower systemic plasma concentrations in females relative to males preferentially decreases the delivered dose of 2,4-D to the proximal tubules in females relative to males.

There were no significant exposure-related effects on other organ weights in P1 and F1 adults (see Supplementary Table S1 for organ weight list). Other histopathological observations in P1 and F1 adults were either isolated findings considered unrelated to 2,4-D exposure and/or spontaneous alterations common to this strain/age of rat.

Weanling body weights and organ weights. Notably, pup assignments at weaning were not performed according to body weight but rather were based on day of birth and litter of origin to control for litter effect and to ensure that each litter is represented in each cohort to the extent possible. PND 22 males in the organ weight subgroup had decreased terminal body weight at all dose levels of 2,4-D (Table 5). Effects on weanling body weight in the 600 ppm group were considered exposure related; however, maternal/pup doses that altered weanling body weights likely exceeded the point of inflection for TK nonlinearity (Saghir *et al.*, 2013). However, no histopathological changes were observed in any organ including the kidney.

Decreased pup body weight at 100 and 300 ppm was considered spurious. Artifactual differences in PND 22 male pup body weights were introduced during group assignment at weaning in the 100 and 300 ppm groups (Table 5). The lack of a dose-response relationship supports a randomization effect. Furthermore, PND 22 weanling body weights in the organ

weight subgroup had a greater body weight differential relative to control animals (9%–10%) at all dose levels than high-dose pups on PND 21 (6%) or high-dose pups randomized into the PND 22 perfusion group (6%). Across all dose levels, body weight decrements on PND 22 were accompanied by decreases in several organ weights, but these organ weight changes did not exhibit consistent dose-response relationships, lacked corresponding histopathology, and were attributed to decreased body weights in this subset. Spearman correlation coefficient confirmed that the weights of the affected organs were highly correlated to body weight ($r \geq .77$) in weanling rats (Table 5).

Together, these results indicated that organ weight changes in male PND 22 weanlings were due to the decreased terminal body weights. Based on the current data set and data in the scientific literature, the sensitivity of organ weights to body weight decrements is age dependent. In a study by Carney *et al.* (2004), a 30% restriction in dietary intake during gestation and lactation resulted in decreased weanling body weights coupled with decreased absolute weights of the liver, kidney, thymus, spleen, uterus, testes, adrenal (males only), and brain (females only). This differs from results in adult male rats, where some organ weights (eg, testes) were conserved in the presence of similar body weight decrements (Chapin *et al.*, 1993).

Reproductive Toxicity

With respect to reproductive toxicity (Supplementary Table S7), 2,4-D had no effects on estrous cyclicity (P1, satellite, Set 3) or P1 reproductive indices, including mating, fertility, time to mating, gestation length, pre- and postimplantation loss, and corpora lutea number (examined in satellite dams). Litter sizes, pup survival, sperm parameters, and ovarian follicle counts

TABLE 5
Weanling Male Body Weights and Organ Weights

PND 22 Organ Weight Subgroup	N	Concentration ^a				Spearman Correlation Coefficient (r)
		0	100	300	600	
Terminal body weight	10 pups/dose	57.9±3.8	52.2±3.8 ^b	52.8±3.9	51.9±7.4^b	NA
Adrenals	10 pups/dose	0.026±0.008	0.022±0.007	0.026±0.010	0.017±0.004^b	0.33
Kidneys	10 pups/dose	0.695±0.085	0.582±0.052 ^b	0.614±0.078	0.589±0.088^b	0.86
Liver	10 pups/dose	2.451±0.322	2.046±0.266 ^b	2.129±0.231	1.998±0.389^b	0.86
Brain	10 pups/dose	1.549±0.091	1.511±0.053	1.492±0.074	1.486±0.070	0.44
Spleen	10 pups/dose	0.282±0.047	0.209±0.045 ^b	0.232±0.028 ^b	0.222±0.050^b	0.77
Testes	10 pups/dose	0.278±0.044	0.237±0.024 ^b	0.239±0.027 ^b	0.236±0.031^b	0.79
Terminal body weights of other subgroups around weaning						
PND 22 perfusion subgroup ^c	12 pups/dose	57.1±3.9	56.4±5.7	52.5±4.8	53.7±6.1	NA
PND 22 combined	22 pups/dose	57.5±3.8	54.5±5.3	52.6±4.3 ^b	52.8±6.6 ^b	NA
PND 21 (all male pups)	24–28 litters/dose	52.5±4.5	51.4±5.4	50.3±4.2	49.4±5.4	NA

Note. Bolded numbers indicate effects interpreted to be exposure related; organ weight decreases were considered secondary to body weight changes.

^aActual dietary concentrations were one-third these values on PND 14–21 due to dietary adjustments to maintain approximately consistent dose levels on a mg/kg/day basis.

^bStatistically different from control mean by Dunnett's test, alpha = .05.

^cBody weights were collected prior to perfusion.

were unaffected by 2,4-D. One P1 male in each of the 300- and 800-ppm groups had decreased bilateral testis size. This incidence of decreased testis size was within the laboratory historical control range (HCD). A similar unilateral finding was noted in a control F1 male. These findings were not reproduced in Set 1a or Set 3 F1 offspring with longer 2,4-D exposures, which included higher mg/kg/day exposures during critical life stages. Furthermore, unilateral and bilateral decreases in testis size have been reported as a spontaneous occurrence in adult Crl:CD(SD) rats (Pettersen *et al.*, 1996). There were no exposure-related histopathological changes in reproductive organs. Overall, there was no indication of reproductive toxicity by 2,4-D in this study.

These results are consistent with the previous 2-generation reproductive toxicity study in F344/DuCrI rats (Rodwell and Brown, unpublished data), which reported no exposure-related effects on reproductive performance at exposures \leq 80 mg/kg/day 2,4-D in the diet. A NOAEL of 5 mg/kg/day was identified in the Rodwell and Brown study because the middose of 20 mg/kg/day produced minimal alterations in postnatal pup body weight that were seen only in the F1b mating group. However, due to a dosing error that elevated actual dose levels only in the F1b mating group, the actual LOAEL in the Rodwell and Brown study was likely higher than 20 mg/kg/day. 2,4-D doses of \geq 80 mg/kg/day produced adverse toxicity in both the P1 animals (decreased body weight and kidney histological alterations) and the pups (decreased pup body weight and survival) (Rodwell and Brown, unpublished data; Saghir *et al.*, 2013), as well as being significantly above the threshold for onset of nonlinear TK (Saghir *et al.*, 2013).

Endocrine Toxicity

Endocrine toxicity includes an evaluation of androgen, estrogen, and thyroid endpoints, as well as other glands that contribute to the maintenance of homeostasis.

Androgenicity/Antiandrogenicity. Androgen-dependent endpoints are summarized in Table 6. In P1 males, decreased seminal vesicle and prostate weights were seen at \geq 300 ppm; prostate weights were not statistically different from those of control. Examination of these data revealed that the absolute and relative organ weights in the control group were atypical, exceeding the laboratory HCD ranges for these organ weights. The organ weights for the 800 ppm-treated males were within the laboratory HCD range (Table 6 footnote). Absolute and relative prostate and seminal vesicle weights also were within HCD values for the 300 ppm-exposed males. Furthermore, there was no associated histopathology in these organs. These findings were not reproduced in Set 1a offspring (PND 70), Set 2b offspring (PND 87–93; data not shown), or Set 3 F1 offspring with longer 2,4-D exposures, which included higher mg/kg/day exposures during critical life stages. There were no alterations in reproductive or accessory sex gland histopathology in Set 1a or Set 3 males. Set 3 males had no effects on

sperm parameters assessed on PND 139. One factor that may have contributed to the statistical significance of seminal vesicle and prostate weights in the P1 males is the inherent variability in these organ weights. Seminal vesicle weights can vary due to differing amounts of seminal fluid, which is influenced by numerous factors (Luke and Coffey, 1994). Inherent variability in prostate weights was recognized in a recent analysis of multigeneration study endpoints (Marty *et al.*, 2009).

There was a slight delay (+ 1.6 days) in F1 preputial separation at 800 ppm, which was attributed to high-dose body weight decrements and slightly delayed growth. Body weight at the time of puberty onset was similar in 800 ppm males and controls (221.1 g compared with 223.2 g in controls). Because high-dose males weighed the same as controls 1.6 days later when puberty onset occurred, these data indicate that 800 ppm 2,4-D had an effect on the rate of growth in peripubescent male rats. The magnitude of this delay (1.6 days with 7%–8% differences in body weights on PND 28–42) was consistent with a 1.8-day delay in age at preputial separation in a feed restriction study with a 10% body weight decrement (Marty *et al.*, 2003). Furthermore, the age at preputial separation in the high-dose males was within the range of historical control values (43.5–44.9 days). Other androgen-sensitive endpoints examined in F1 offspring, including AGD and nipple retention, which are considered very sensitive endpoints (Clark, 1999; Hotchkiss *et al.*, 2004; McIntyre *et al.*, 2001; Wolf *et al.*, 2002), male reproductive organ weights, sperm parameters, and testicular and accessory sex gland histopathology, were not altered. In the previous 2-generation toxicity study, there were no exposure-related alterations in testicular (P1, F1b), epididymal (P1, F1b), prostate (F1b), or seminal vesicle (F1b) histopathology with 2,4-D exposure (Rodwell and Brown, unpublished data); the F1b treatment was \geq 100 mg/kg/day, substantially higher than the top dose in this study. Last, 2,4-D TK was nonlinear in F1 males at 800 ppm, which casts further doubt on the toxicological significance of any effects to human risk at this dose level.

With redundancy in several endpoints examined across generations, a weight-of-evidence approach was applied to discern consistent effects attributable to 2,4-D treatment. Overall, the data do not support an androgenic or antiandrogenic potential for 2,4-D, even at doses exceeding the threshold for linear TK in rats.

Estrogenicity/Antiestrogenicity. Data for estrogen-sensitive endpoints, which included estrous cyclicity, reproductive indices, organ weights/pathology, and ovarian follicle counts, are shown in Table 7. Reproductive indices and litter size and pup survival were not affected by 2,4-D. There were no exposure-related effects on age and body weight at vaginal opening (measured in all Sets 1–3 animals). There were no effects on female reproductive organ histopathology in either P1 or F1 offspring, or quantitative ovarian follicle counts in F1 Set 3 offspring. There were no significant, exposure-related changes in

TABLE 6
Androgen-Dependent Endpoints Evaluated in the 2,4-D EOGRTS

	0 ppm	100 ppm	300 ppm	800 ppm	0 ppm	100 ppm	300 ppm	800 ppm				
Parameter	P1 Males				F1 Males/F1 Females ^a							
AGD (mm) (M/F)	—	—	—	—	3.92/2.08	3.78/2.03	3.83/2.06	3.74/2.01				
AGD (mm) ³ /body weight) (M/F)	—	—	—	—	2.02/1.09	1.95/1.06	1.97/1.08	1.95/1.06				
	P1 Males				F1 Males ^a							
Nipple retention on PND 13	—	—	—	—	0.0	0.0	0.0	0.0				
Age at preputial separation (PPS)	—	—	—	—	43.0	43.6	43.3	44.6 ^b				
Body weight at PPS (g)	—	—	—	—	223.2	230.3	221.9	221.1				
	P1 Males ^c				Set 3 Males							
Sperm motility (% motile)	97.9	97.7	94.6	94.7	96.8	97.4	97.4	97.4				
Sperm motility (% progressively motile)	89.0	86.9	85.2	85.4	88.0	86.5	86.6	87.4				
Spermatid counts (×10 ⁶)	222.8	—	—	238.9	344.2	—	—	322.5				
Spermatid counts (Conc/g testis) × 10 ⁶	119.7	—	—	135.1	184.7	—	—	176.2				
Sperm counts (×10 ⁶)	235.3	—	—	226.3	278.5	—	—	292.7				
Sperm counts (Conc/g epididymis) ×10 ⁶	669.2	—	—	681.1	833.7	—	—	895.6				
Proportion abnormal sperm	0.014	—	—	0.020	0.016	—	—	0.025				
	P1 Males ^d				Set 1a Males ^d				Set 3 Males ^d			
	0 ppm	100 ppm	300 ppm	800 ppm	0 ppm	100 ppm	300 ppm	800 ppm	0 ppm	100 ppm	300 ppm	800 ppm
Absolute prostate (g)	1.313	1.274	1.198	1.188	0.753	0.776	0.791	0.707	1.041	1.075	1.081	1.070
Relative prostate (g/100 g body weight)	0.248	0.242	0.228	0.221	0.198	0.207	0.209	0.208	0.181	0.174	0.186	0.187
Absolute testes (g)	3.718	3.654	3.595	3.528	3.306	3.366	3.201	3.237	3.743	3.801	3.684	3.725
Relative testes (g/100 g body weight)	0.706	0.696	0.686	0.653	0.875	0.904	0.849	0.954	0.651	0.620	0.636	0.645
Absolute seminal vesicles (g)	1.887	1.811	1.653 ^b	1.655 ^b	1.072	1.147	1.141	1.010	1.807	1.741	1.760	1.761
Relative seminal vesicles (g/100 g body weight)	0.358	0.347	0.315 ^b	0.308 ^b	0.282	0.307	0.301	0.298	0.313	0.284	0.304	0.308
Absolute epididymides (g)	1.498	1.428	1.405	1.418	0.895	0.862	0.870	0.842	1.399	1.386	1.375	1.388
Relative epididymides (g/100 g body weight)	0.285	0.272	0.268	0.263	0.237	0.232	0.232	0.248	0.244	0.226	0.237	0.243
Terminal body weight (g)	529.7	528.8	530.6	540.7	379.4	373.2	379.3	340.1 ^b	578.7	616.1 ^b	585.3	577.4

Note. Italicized numbers indicate control values.

^an = 24–28 litters/dose level.

^bStatistically different from control mean by Dunnett's test, alpha = .05.

^cn = 27 per dose level for P1 sperm parameters; n = 23–27 per dose level for Set 3 sperm parameters.

^dn = 27 per dose level for P1 organ weights; n = 10 per dose level for Set 1a organ weights; n = 23–27 per dose level for Set 3 organ weights.

reproductive organ weights in P1 or F1 PND 22 females or in F1 Set 1a, Set 2 (testis), or Set 3 males and females. When uterine weights were examined in P1 and F1 females, there was a suggestion of increased uterine weight in the high-dose group; however, this result was deemed incidental and unrelated to 2,4-D treatment. In each case, the increase in uterine weights was not statistically significant (10%–17% in P1 and Set 3 [n = 24–27]; 31% in Set 1a [n = 10]) and showed high variability (eg, coefficients of variation were 21%–36%) because

stage of estrous was not controlled at necropsy. An evaluation of terminal stage of estrous indicated that the increased uterine weights generally correlated with a higher incidence of 600 ppm females in proestrus or estrus at the time of necropsy. Given that 2,4-D did not alter estrous cycle patterns in either P1 or F1 Set 3 offspring (no difference in percent of days in estrus or diestrus) and there were no differences in mean cycle length (4.1 days in all P1 groups, and 4.4 vs 4.3 days in F1 Set 3 control and high-dose groups, respectively), the increased

TABLE 7
Estrogenic-Dependent Endpoints Evaluated in the 2,4-D EOGRTS

	0 ppm	100 ppm	300 ppm	600 ppm	0 ppm	100 ppm	300 ppm	600 ppm				
Parameter	P1 Females				F1 Females ^a							
Age at vaginal opening (VO)	—	—	—	—	32.0	31.9	31.8	32.3				
Body weight at VO (g)	—	—	—	—	108.8	106.5	106.0	104.1				
	P1 Females (Main and Satellite) ^b				Set 3 Females ^b							
Mean estrous cycle length	4.1	4.1	4.1	4.1	4.4	4.2	4.4	4.3				
Ovarian follicle counts (small)	—	—	—	—	66	—	—	66				
Ovarian follicle counts (growing)	—	—	—	—	25	—	—	28				
	P1 Females ^c				Set 1a Females ^c				Set 3 Females ^c			
	0 ppm	100 ppm	300 ppm	600 ppm	0 ppm	100 ppm	300 ppm	600 ppm	0 ppm	100 ppm	300 ppm	600 ppm
Absolute uterus (g)	0.677	0.668	0.657	0.793	0.5229	0.5354	0.5521	0.6861	0.679	0.734	0.750	0.750
Relative uterus (g/100 g body weight)	0.229	0.225	0.219	0.269	0.2307	0.2351	0.2436	0.3039	0.228	0.245	0.250	0.252
Absolute ovaries (g)	0.113	0.118	0.119	0.111	0.099	0.105	0.105	0.108	0.089	0.096	0.094	0.091
Relative ovaries (g/100 g body weight)	0.038	0.040	0.039	0.037	0.044	0.046	0.045	0.048	0.030	0.032	0.031	0.031
No. estrus or proestrus at necropsy (No. affected/No. evaluated)	5/21 ^{d,e}	6/21 ^{d,e}	4/21 ^d	6/21 ^d	3/10	3/10	3/10	6/10	10/27	6/27	11/23	5/24
Terminal body weight (g)	295.8	297.5	300.5	296.5	229.6	228.3	230.9	224.6	299.5	299.0	304.3	300.0

Note. Italicized numbers indicate control values.

^an = 24–28 litters/dose level.

^bn = 39 per dose level for P1 estrous cycle length; n = 23–27 per dose level for Set 3 estrous cycle length and n = 15 per dose level for ovarian follicle counts.

^cn = 24–28 per dose level for P1 organ weights; n = 10 per dose level for Set 1a organ weights; n = 23–27 per dose level for Set 3 organ weights.

^dTerminal estrus inadvertently not determined in 7, 7, 6, and 6 P1 animals in the 0, 100, 300, and 600 ppm groups, respectively.

^eIncluded 1 additional animal per dose group transferred to the main study from the satellite group.

number of animals in proestrus/estrus at necropsy appears to be incidental. Examination of HCD revealed that control uterine weights (P1 dams on LD 22, Set 1a on PND 70, Set 3 on PND 139) were below the laboratory HCD range, whereas high-dose absolute and/or relative uterine weights were within the laboratory HCD range. These findings further support the lack of an exposure-related effect on uterine weights.

The absence of effects on the ovary and uterus is consistent with the previous 2-generation toxicity study (Rodwell and Brown, unpublished data), which reported no exposure-related alterations in ovarian (P1, F1b) or uterine (P1 adults; F1b adults and weanlings; F2a weanlings) histopathology with 2,4-D exposure (estrous cyclicity and vaginal opening were not assessed by Rodwell and Brown). Notably, the F1b and F2a weanlings were not cycling, so the absence of gross findings (eg, enlargement or histopathology) gives a more robust assessment of potential estrogenic effects than looking at uterine weights in cycling females. Thus, the data do not support any 2,4-D-mediated estrogenic or antiestrogenic effects, even at doses exceeding the threshold for linear TK in rats.

Thyroid assessment. Thyroid assessment included thyroid hormones (T3, T4, and TSH), thyroid weights, and/or histopathology evaluated at multiple life stages (Table 8). There was no

consistent pattern of effects on thyroid parameters seen across life stages. High-dose satellite GD 17 dams given 600 ppm 2,4-D had nonsignificant decreases in T4 and T3 with a corresponding increase in TSH. Thyroid histopathological alterations were seen in 3 of 12 dams, which comprised smaller thyroid follicles with small vacuoles in the colloid that were suggestive of colloid resorption. There were no adverse pathological alterations (eg, degeneration) associated with the smaller follicles, and thyroid changes were not observed in LD 21 dams, indicating that this effect was transient; thus, these thyroid changes were considered adaptive. High-dose alterations in thyroid function during pregnancy is plausible given that high doses of 2,4-D have been shown to compete with T4 for serum protein binding (Florsheim and Velcoff, 1962; Florsheim *et al.*, 1963; Van den Berg *et al.*, 1991) and because pregnancy affects all aspects of thyroid hormone economy (Larsen and Ingbar, 1992). This adaptive change in thyroid function in GD 17 dams occurred only at 600 ppm, an exposure level that resulted in nonlinear TK (particularly exacerbated in GD 17 dams compared with nonpregnant adult females); thus, this finding was considered not relevant for human risk assessment. Furthermore, rats are a more sensitive model to thyroid perturbations than humans (Jahnke *et al.*, 2004; McClain, 1995), with T4 less tightly bound to serum protein in rats than in humans (Jahnke *et al.*,

TABLE 8
Thyroid-Related Endpoints Evaluated in the 2,4-D EOGRTS

Life Stage	PPM	N	Male				Female				Male and Female Thyroid Histopath ^c
			T ₃ (ng/dl)	T ₄ (µg/dl)	TSH ^b (ng/ml)	Thyroid weight ^b (g/100 g body weight)	T ₃ (ng/dl)	T ₄ (µg/dl)	TSH ^b (ng/ml)	Thyroid weight ^b (g/100 g body weight)	
P1 males	0	27-28	—	—	—	0.0045	—	—	—	0.0061	No treatment-related effects
	100	27-28	—	—	—	0.0045	—	—	—	0.0057	
	300	27	—	—	—	0.0044	—	—	—	0.0054 ^d	
	800	27	—	—	—	0.0047	—	—	—	0.0058	
GD 17 dams	0	11	—	—	—	—	73.12	1.26	2.92	0.0043	3 High-dose dams: very slight decrease in follicle size
	100	10	—	—	—	—	71.69	1.22	2.76	0.0048	
	300	10	—	—	—	—	69.64	1.16	2.60	0.0043	
	600	12	—	—	—	—	68.12 ^e	1.15 ^e	3.65 ^e	0.0047	
PND 4: Cull pups	0	9-10	34.51	0.64	1.12	—	41.99	0.85	0.97	—	—
	100	7-10	35.90	0.62	0.98	—	38.64	0.99	1.06	—	
	300	6-10	35.46	0.55	1.06	—	36.59	0.72	0.90	—	
	600	9-10	32.19	0.56	1.09	—	40.29	0.73	1.15	—	
PND 22: F1 unselected weanlings	0	10	107.22	3.62	1.32	0.0109	99.14	3.57	0.99	0.0112	No treatment-related effects
	100	10	100.82	4.40	1.25	0.0104	110.43	3.99	1.13	0.0109	
	300	10	86.56 ^d	2.98	1.48	0.0096	99.42	3.55	0.94	0.0100	
	600	10	93.46	2.59 ^d	1.27	0.0105	107.42	2.85	1.02	0.0108	
PND 62-64: F1 Set 1a	0	10	78.69	4.75	2.95	0.0051	67.08	2.35	1.89	0.0070	No treatment-related effects
	100	10	69.78	4.46	3.21	0.0055	66.89	2.27	2.05	0.0069	
	300	10	66.77	5.31	3.72	0.0050	70.45	2.80	2.10	0.0063	
	600/800	10	72.03	4.11	3.62	0.0052	74.28	2.79	2.34	0.0067	
PND 138-140: F1 Set 3	0	27	—	—	—	0.0046	—	—	—	0.0059	No treatment-related effects
	100	27	—	—	—	0.0045	—	—	—	0.0064	
	300	23	—	—	—	0.0045	—	—	—	0.0062	
	600/800	23-24	—	—	—	0.0045	—	—	—	0.0063	

Note. Italicized numbers indicate control values. Bolded values and findings are considered exposure related; —, not applicable.

^aFor hormone measurements, $n = 10-12$ per dose level for GD 17 dams; 6-10 per dose level for PND 4 pups; $n = 10$ per dose level for PND 22 and Set 1a offspring.

^bFor thyroid weights, $n = 24-28$ /sex/dose level for P1 thyroid weights; $n = 10-12$ per dose level for GD 17 dams; $n = 9-10$ /sex/dose level for PND 22 weanlings; $n = 10$ /sex/dose level for Set 1a; $n = 23-27$ /sex/dose level for Set 3.

^c $n = 24-28$ per dose level for P1 organ weights; $n = 11-12$ per dose level for GD 17 dams; $n = 9-10$ /sex/dose level for PND 22 weanlings; $n = 10$ /sex/dose level for Set 1a; $n = 10$ /sex/dose level for Set 3.

^dStatistically different from control mean by Dunnett's test, $\alpha = 0.05$.

^eHormone values were not statistically identified, but changes considered possibly exposure related due to the pattern of findings and associated adaptive histopathology.

2004). There were no thyroid effects at lower dose levels in GD 17 dams and no biologically significant effects on thyroid endpoints at the other life stages examined.

Adrenal and pituitary. There was no indication of alterations in adrenal function as there were no effects on absolute or relative adrenal gland weights and no exposure-related histopathological findings at any 2,4-D exposure level. Absolute and relative (fixed) pituitary gland weights were significantly decreased in the 800 ppm Set 3 males; however, this finding was not considered exposure related. The magnitude of the differences from pituitary weights in control animals was minimal (absolute and relative pituitary weights were 0.0127 and 0.0022 g/100 g body weight, respectively, compared with 0.0139 and 0.0024 g/100 g body weight in controls), and values were within the HCD for this laboratory (data not shown). Pituitary weights in P1 and Set 1a males, as well as in P1, Set 1a, or Set 3 females, did not differ from control values, and there were no exposure-related histopathological changes in pituitaries at any dose level or life stage. Last, alterations in pituitary function would be anticipated to alter numerous study endpoints in Set 3 males, including reproductive and accessory sex gland weights and sperm parameters, and these endpoints were not affected by 2,4-D exposure. Thus, despite the pituitary weight differences from control in the Set 3 males, there is no indication of pituitary alterations with 2,4-D exposure.

Triggers for Second-Generation Mating

In the EOGRTS design, the decision to breed a second generation is triggered based on predetermined criteria. With the lack of significant treatment-related effects on reproductive parameters and developmental landmarks, mating of a second generation was not triggered based on *a priori* established criteria (Table 9). The triggers required that a statistically significant, or biologically relevant, dose-related response was demonstrated. As with other toxicological studies, weight of evidence was applied when interpreting the results of the EOGRTS to determine whether a second-generation breeding was required.

The laboratory's HCD also assisted in the interpretation of data on reproductive toxicity endpoints. MOE considerations advocated by the ACSA program also were factored into the decision not to breed a second generation (Cooper *et al.*, 2006). The ACSA paradigm indicated that if toxicity triggers were limited to the high-dose level alone (with no apparent dose-related trend), the MOE of this dose relative to either estimated human exposures or those directly measured through human biomonitoring studies should be considered in the triggering decision. If the MOE was judged sufficient, triggering of a second-generation breeding would not be required. The MOE for 2,4-D exposure, based on data from the Farm Families Exposure Study (Alexander *et al.*, 2007), indicated that children were exposed to geometric mean doses of 2,4-D (0.22 µg/kg/day) that were more than 22 000-fold lower than the projected NOEL of 5 mg/kg/day for the EOGRTS. Using the estimated internal dose in pups in the range-finding/TK study (Saghir *et al.*, 2013), the MOE of exposed pups was > 13 000-fold higher than estimated children's systemic exposures. Based on these results, breeding of a second generation was not triggered in this EOGRTS.

DNT

Neurobehavioral and neuropathological endpoints also were examined in this EOGRTS study. There were no clinical observations related to 2,4-D treatment in any dose group. Potential DNT effects were examined in Set 1b F1 offspring exposed to 2,4-D from PND 21–60. FOB parameters (handheld and open-field observations, grip performance, landing foot splay, and rectal temperature), MA, and ASR were assessed in these animals on PND 54–55 (results shown in Supplementary Figures S6–S12). On PND 60, these F1 adults were perfused for central and peripheral nervous system neuropathology. In addition, brain weights, gross brain measurements, and morphometric measurements were evaluated. A subset of unselected weanlings also was perfused for neuropathology, brain weights, and gross brain measurements. There were no significant, exposure-related effects on FOB parameters, MA, or ASR. Gross and morphometric brain measurements also were not affected.

TABLE 9
Triggers for Breeding a Second Generation

Endpoint	Comment	Trigger Met?
P1 estrous cycle evaluation	Biologically relevant and dose related without overt toxicity	No
P1 fertility	In the absence of reproductive organ histopathology	No
F1 litter parameters	Biologically relevant and dose related without severe maternal toxicity	No
F1 developmental landmarks (AGD)	In the absence of body weight-mediated changes	No
F1 developmental landmarks (nipple retention)	In the absence of body weight-mediated changes	No
F1 developmental landmarks (male puberty onset)	In the absence of body weight-mediated changes	Delayed in the presence of decreased body weight
F1 developmental landmarks (female puberty onset)	In the absence of body weight-mediated changes	No
F1 estrous cycle evaluation	Biologically relevant and dose related without overt toxicity	No

There was no exposure-related neuropathology in PND 22 or 60 animals, including no effects on brain myelin deposition (assessed by special staining).

Earlier studies using 2,4-D doses of 70–100 mg/kg/day administered either IP or SC produced evidence of altered myelin deposition (Duffard *et al.*, 1996; Rosso *et al.*, 2000). In contrast, in this study, 2,4-D had no effect on myelin deposition despite treatment during sensitive periods of postnatal myelin deposition in rats and use of doses that were well above the saturation threshold for renal clearance.

Related neurotoxicity endpoints were examined in P1 adults and in other F1 offspring, including brain weights and/or histopathology in nonperfused animals (P1 adults and PND 22, Set 1a on PND 70, and Set 3 on PND 139). Terminal body weights and absolute brain weights were significantly decreased in PND 70 high-dose males; however, this group had a smaller sample size ($n = 10/\text{sex}/\text{dose}$) and shorter exposure duration than other groups with no brain weight changes (P1 adults and F1 Set 3 on PND 139; both with $n \geq 23/\text{sex}/\text{dose}$); thus, the brain weight finding at PND 70 is considered spurious. There were no exposure-related effects on brain histopathology at any age. Thus, there was no evidence of DNT related to 2,4-D exposure.

DIT

Potential DIT of 2,4-D was evaluated in an EOGRS (Supplementary Figures S13–S14). To examine TDAR function, Set 2a AFC-designated animals were immunized (IV) with 2×10^8 SRBC on PND 66–70. After 4 days, spleen cells were evaluated for AFC responses. 2,4-D had no effect on AFC responses in males. High-dose females had nonsignificant decreases in AFC/spleen and AFC/ 10^6 splenocytes (54% and 27%, respectively). The AFC response was not altered at ≤ 300 ppm in females. The observation of lower AFC response in 600 ppm Set 2a females appeared to be due to temporal variability over the 4-day span of the evaluations, although the variability was within the HCD range for the laboratory and the range reported in the literature for this assay (Mann *et al.*, 2008; White *et al.*, 2005). Analysis of the data by day reveals that days 1 and 2 of the assay yielded a stronger AFC response compared with days 3 and 4 and, unfortunately, dose groups were not balanced across days due to predefined age restrictions for the assay and the need to complete the AFC assay with the first half of the available litters to determine whether the NK assay was required. In the EOGRS test guideline (OECD 443), only the TDAR assay is recommended for DIT assessments to alleviate this issue. The lack of effect of 2,4-D on the primary immune response to SRBC is consistent with previously published reports in mice in which 2,4-D did not suppress the primary antibody response to SRBC (Blakley, 1986; Blakley and Blakley, 1986; Blakley and Schiefer, 1986; Lee *et al.*, 2001). Due to the contribution of interexperimental variability to the SRBC decrease, the presence of a lower SRBC response only at a dose that exhibits nonlinear TK (600 ppm), and previous published studies that demonstrate a lack of effect of 2,4-D on the

SRBC response, it was determined that 2,4-D did not induce an exposure-related decrease in the primary immune response to SRBC in male or female rats; consequently, the NK-cell assay was conducted in the Set 2b animals.

For the F1 NK assay on PND 87–93 (Set 2b), fluorescent labeled YAC-1 cells (targets, T) were plated with spleen cells (effectors, E) at E:T ratios from 50:1 to 800:1. Target cytotoxicity was monitored directly through PI labeling using flow cytometry. The NK assay showed no effects from 2,4-D exposure. The assay showed linear cytotoxicity with increasing E:T cell ratios (from 5% to 40% cytotoxicity) which was identical across all doses.

There were no exposure-related alterations in other immune-related endpoints in P1 or F1 animals, which support the conclusion of no relevant changes in immune system development or function. There were no exposure-related effects on spleen and thymus weights (P1, Set 1a, Set 2a, Set 2b [spleen only] and Set 3) or histopathological changes in any of the immune-related tissues examined (ie, spleen, mediastinal lymph nodes, mesenteric lymph nodes, spleen and/or thymus in the P1, Set 1a, and Set 3 animals). White blood cell parameters were not altered in P1 or Set 1a animals, the 2 groups for which hematological analyses were conducted.

Thus, it was concluded that 2,4-D did not exhibit immunosuppressive properties in either the SRBC AFC assay or the NK-cell assay. The combination of these assays has previously been reported to accurately identify immunotoxic compounds with a 94% concordance (Luster *et al.*, 1992). Therefore, the results of these assays indicate that 2,4-D is not immunotoxic. In addition, because 2,4-D treatment was continuous throughout all life stages up to adult in which immunotoxicity was assessed, these findings indicate 2,4-D is not a developmental or adult immunotoxicant.

Advantages of the EOGRS Study Design

The EOGRS study provides a thorough assessment of reproductive toxicity, DNT, DIT, endocrine toxicity, and systemic toxicity in animal exposed during critical windows of development (ie, F1 animals are exposed during gestation, lactation and through adulthood). The assignments into cohorts may be altered if concern regarding a particular type of effect predominates based on other toxicity data. A comparison of the 2,4-D EOGRS study, the recently adopted EOGRS test guideline (OECD 443) and the test guidelines/study guidance documents for the 2-generation reproductive toxicity study, DNT study, DIT study (proposed design) and comparative thyroid study are provided in Supplementary Tables S8–S10.

The EOGRS provides the opportunity to evaluate multiple systems for toxicity in the F1 offspring exposed during critical windows of development. These data can be used to determine which target organs are most sensitive to toxicant-induced effects. Furthermore, this design provides an opportunity to evaluate toxicity across systems (eg, did renal toxicity contribute to alterations in neurobehavioral performance? Did

endocrine toxicity occur at doses lower or higher than doses affecting systemic toxicity endpoints?).

The EOGRTS study also provides the opportunity to evaluate the same or related endpoints at multiple time points to examine whether there are reproducible patterns of effects (eg, are effects on P1 male reproductive organ weights reproducible in F1 offspring exposed to greater concentrations of test material for longer periods of time, including sensitive periods of development and maturation?). With so many endpoints evaluated in an EOGRTS study and a 5% chance for Type II errors ($\alpha \leq .05$), there is a likelihood that some statistically identified values are not toxicologically meaningful; therefore, intra-study reproducibility of these results across F1 cohorts can be useful to establish treatment-related findings. If equivocal data occur during conduct of the EOGRTS, the use of multiple cohorts (sets) often provides an opportunity to evaluate endpoints in additional animals (eg, inclusion of testes weights in the DIT cohort).

Another advantage of the EOGRTS design is the recommendation to use TK data for dose setting and data interpretation. Use of TK data in this manner is predicated on a sufficient MOE in humans to verify that humans are unlikely to exhibit nonlinear TK when exposed to the test material. As was demonstrated in this study, this paradigm avoids use of unrealistically high doses that may produce toxicities that are ultimately irrelevant to human risk, and instead, allows the selection of dose levels that generate dose-response data that are more relevant for risk assessment purposes. Collection of TK data under conditions of which the compound is being tested and are most relevant to real-world human exposure, ie, dietary as exemplified by this study, also allows for both a more informed study design and interpretation of toxicological findings (eg, maternal and neonate internal dosimetry changes between life stages and the need for dose adjustment) (Saghir *et al.*, 2013).

One major advantage of the EOGRTS study is that it gathers more information while using significantly fewer animals. When the F2 breeding of a second generation is not required based on a failure to fulfill the pre-established trigger criteria, the EOGRTS study uses approximately 1600 animals as compared to approximately 1000 additional animals if an F1 mating is triggered (assumes 4 dose groups of 20 litters each and 13 pups/litter from F1 breeding). Even more importantly, however, approximately 6000 animals would have been used if the reproductive, DNT, DIT and comparative thyroid endpoints evaluated for 2,4-D in the EOGRTS had been evaluated in stand-alone test protocols (assumes 10 pregnant dams/dose with 4 dose groups in the DIT). Furthermore, the EOGRTS study is considered a Tier 2 study for endocrine testing, thereby obviating the need to conduct the mammalian assays in the U.S. EPA Endocrine Disruptor Screening Battery (U.S. EPA, 2013). An associated animal welfare benefit is that the KMD-based dose selection strategy encouraged in the EOGRTS study design provides a toxicologically-defensible rationale for reducing unnecessary stress and/or discomfort to test animals

that might otherwise be encountered in an MTD-based dose selection strategy.

The EOGRTS study also examines 3–4 pups/sex/litter as opposed to 1 pup/sex/litter in the 2-generation reproductive toxicity study, which improves the probability of detecting low incidence events. Thus, a more thorough toxicity assessment is conducted while minimizing animal usage to the extent possible. Lastly, the EOGRTS study has a shorter in-life phase than the 2-generation reproductive toxicity study, lasting approximately 25 weeks when a second generation is not bred compared to approximately 36 weeks for a 2-generation study.

Caveats for the 2,4-D Study Design

The EOGRTS study is a relatively complex study design and is resource intensive. Generally, because of staffing requirements, the EOGRTS is more expensive than a 2-generation reproductive toxicity study, but less expensive than separate reproductive toxicity and DNT studies. Laboratories that have experience conducting the 2-generation reproductive toxicity study, DNT study, immunotoxicity studies and endocrine studies have the experience needed to conduct the EOGRTS, but should carefully plan for the logistical challenges of managing the different F1 cohorts. Particularly challenging is the time around weaning, when P1 lactating dams and PND 22 weanlings are necropsied and other F1 animals are assigned to cohorts. These study phases require careful planning to determine pup assignments prior to weaning and assigned staff to manage the number of animals requiring various assessments. If appropriate, the study can be conducted in replicates to facilitate animal management (eg, 2 or 3 blocks with staggered starts 1 week apart).

To meet the logistical demands around weaning, an *a priori* plan is needed to assign pups to cohorts. Pups born on different days are assigned to cohorts when they reach weaning, which does not allow for the randomization of pups by body weight at weaning. This can result in body weight differences across cohorts at this life stage. Artifactual differences in weanling body weights must be considered when examining endpoints such as organ weights and reproductive maturation. If there is a true effect on body weights, a consistent pattern is likely across cohorts.

Unless triggered to breed a second generation, the EOGRTS study examines reproductive behavior in the P1 animals only. Although only limited data are collected from the production of the second generation offspring, a second mating allows for a replicate assessment of reproductive performance in the F1 animals, as well as F2 litter size, offspring survival and development (including AGD and nipple retention), and weanling necropsy endpoints (organ weights and histopathology). However, if effects on neurotoxicity or immunotoxicity are seen in the first generation, these parameters are not assessed in the F2 (second generation) offspring in the standard EOGRTS design. Thus, if these endpoints are sensitive to toxicant effects, breeding of the second generation offers no advantage for risk

assessment; instead, a follow-up study to more fully characterize DNT or immunotoxicity would be more beneficial. With the exception of puberty onset, many adult endpoints are examined in both the P1 and F1 offspring, including estrous cyclicity, reproductive organ weights and histopathology, sperm analyses (motility, counts and morphology) and qualitative/quantitative ovarian examinations. With males, it is generally accepted that reproductive organ histopathology and sperm parameters are more sensitive endpoints than fertility for detecting changes in the male reproductive system (Chapin *et al.*, 1997; Gray *et al.*, 1989; Mangelsdorf *et al.*, 2003; Ulbrich and Palmer, 1995). Functional evaluations of fertility are less sensitive due to the excess sperm reserve in rodents. Thus, while P1 males are not exposed for a full spermatogenic cycle (10 weeks) prior to breeding, males are exposed for ≥ 10 weeks prior to necropsy, when endpoints more sensitive than fertility are assessed. Female rats have robust reproductive performance and hormone production even in the presence of reproductive system alterations (eg, within 24h of removing an ovary, the remaining ovary can ovulate a full complement of 10–12 follicles; also only 2–3 corpora lutea are needed to maintain pregnancy; Hirshfield, 1987). Thus, ovarian histopathology is believed to be a sensitive indicator of female reproductive toxicity (Regan *et al.*, 2005). Female reproductive organ weights, however, are highly variable depending on the stage of the estrous cycle at necropsy and are not sensitive endpoints in this study design.

While some endpoints are only evaluated in F1 offspring (eg, puberty onset), the evaluation of 3–4 pups/sex/litter ensures greater precision in these data and a greater likelihood of finding an effect if one exists due to improved statistical power (OECD, 2013). Thus, a second generation is bred only when *a priori* criteria are met (see OECD Guidance Document 117; OECD, 2011). This approach has been supported by several retrospective analyses of 2-generation reproductive toxicity studies, including an evaluation of over 500 studies (Janer *et al.*, 2007; Piersma *et al.*, 2011; Rorije *et al.*, 2011).

Another caveat in using the EOGRTS design is that it requires timely data analysis to determine if a breeding trigger for the F2 generation has been met. Laboratories must identify deadlines for data collection and analysis, and designate resources to complete these evaluations in a timely manner, to ensure data are available to evaluate triggers for breeding a second generation. For studies implemented for regulatory compliance, participation in this decision by the relevant regulatory authorities will also increase confidence in the decision; this however requires keeping these authorities informed of the study designs and results on a real time basis.

Most endpoints in the EOGRTS design are collected by age; thus, the number of animals available for data collection (and the number of times an experimental procedure must be conducted) depends on the distribution of birth dates. For some endpoints (eg, SRBC assay), inter-assay variability requires that the number of assays be limited to the extent possible. For these endpoints, it is important to allow a range of ages over which

data can be collected and recognize that some endpoints (eg, spleen weight) may be affected by allowing use of an age range.

Overall, the EOGRTS provided a robust assessment of potential reproductive, endocrine, neurodevelopmental, immune and systemic toxicity for 2,4-D in animals exposed during critical windows of development. Consistent with previous toxicity studies, renal toxicity appears to be the most sensitive target organ following 2,4-D exposure. TK data played an integral role in this study design with the selection of 2,4-D doses both below and above the threshold for nonlinear TK and the use of TK data to interpret toxicity findings. Aside from renal effects, 2,4-D did not alter reproductive, endocrine, neurodevelopmental, or immune endpoints at doses within the linear TK range, even with exposures during critical stages of development.

SUMMARY

Across life stages, the EOGRTS has established that 2,4-D showed no evidence of adverse effects on reproduction, DNT or DIT; thus, the NOAEL for each of these toxicities was 800 ppm in males and 600 ppm in females, the highest dose levels tested. There was no evidence of potential interactions with the androgen or estrogen endocrine pathways, or of interference with the steroidogenesis pathway or hypothalamus-pituitary-gonadal axis. In addition, separate TK characterizations indicated that renal clearance of 2,4-D was saturated in males at 800 ppm and females at 300 ppm. Effects on the thyroid were limited to adaptive changes during gestation at a dose exceeding the linear saturation threshold, and thus irrelevant for human health risk assessment. Thus, the NOAEL for endocrine toxicity (estrogen, androgen, and thyroid pathways) also was 800 ppm in males and 600 ppm in females. Kidney was a primary target organ with very slight to slight degeneration of proximal convoluted tubules observed in high-dose P1 males and high-dose F1 males and females. Due to these findings, the NOAEL for systemic toxicity was 300 ppm in both males (16.6 mg/kg/day) and females (20.6 mg/kg/day). In addition, human exposures to 2,4-D have been measured for adults and children on farms actively using 2,4-D (geometric mean doses of 2.46, 0.8, and 0.22 $\mu\text{g/kg}$ for applicators, spouses, and children, respectively; Alexander *et al.*, 2007). These exposures are $> 6000\times$ lower than the identified NOAELs in the 2,4-D EOGRTS study; therefore, these data indicate that 2,4-D is unlikely to pose significant health risks to humans.

SUPPLEMENTARY DATA

Supplementary data are available online at <http://toxsci.oxfordjournals.org/>.

FUNDING

Industry Task Force II on 2,4-D Research Data.

ACKNOWLEDGMENTS

Authors would like to acknowledge Darlene Curell, Rebecca Drury, Kathy Gallagher, Barb Hall, Jane Lacher, Marie Lawson, Lisa McFadden, Beth Stieve, Amy Clark, Jennifer Staley, Barry Yano, and the histology and animal care staff of The Dow Chemical Company for their help in completing this study. The authors work for a company at which 2,4-D is produced or consult for the Industry Task Force II on 2,4-D Research Data.

REFERENCES

- Alexander, B. H., Mandel, J. S., Baker, B. A., Burns, C. J., Bartels, M. J., Acquavella, J. F., and Gustin, C. (2007). Biomonitoring of 2,4-dichlorophenoxyacetic acid exposure and dose in farm families. *Environ. Health Perspect.* **115**, 370–376.
- Aylward, L. L., Morgan, M. K., Arbuckle, T. E., Barr, D. B., Burns, C. J., Alexander, B. H., and Hays, S. M. (2010). Biomonitoring data for 2,4-dichlorophenoxyacetic acid in the United States and Canada: Interpretation in a public health risk assessment context using Biomonitoring Equivalents. *Environ. Health Perspect.* **118**, 177–181.
- Barton, H. A., Pastoor, T. P., Baetcke, K., Chambers, J. E., Diliberto, J., Doerrer, N. G., Driver, J. H., Hastings, C. E., Iyengar, S., Krieger, R., et al. (2006). The acquisition and application of absorption, distribution, metabolism, and excretion (ADME) data in agricultural chemical safety assessments. *Crit. Rev. Toxicol.* **36**, 9–35.
- Beringer, P., Nguyen, M., Hoem, N., Louie, S., Gill, M., Gurevitch, M., and Wong-Beringer, A. (2005). Absolute bioavailability and pharmacokinetics of linezolid in hospitalized patients given enteral feedings. *Antimicrob. Agents Chemother.* **49**, 3676–3681.
- Blakley, B. R. (1986). The effect of oral exposure to the n-butylester of 2,4-dichlorophenoxyacetic acid on the immune response in mice. *Int. J. Immunopharmacol.* **8**, 93–99.
- Blakley, B. R., and Blakley, P. M. (1986). The effect of prenatal exposure to the n-butylester of 2,4-dichlorophenoxyacetic acid (2,4-D) on the immune response in mice. *Teratology* **33**, 15–20.
- Blakley, B. R., and Schiefer, B. H. (1986). The effect of topically applied n-butylester of 2,4-dichlorophenoxyacetic acid on the immune response in mice. *J. Appl. Toxicol.* **6**, 291–295.
- Boelen, A., Wiersinga, W. M., and Fliers, E. (2008). Fasting-induced changes in the hypothalamus-pituitary-thyroid axis. *Thyroid* **18**, 123–129.
- Buist, S. C., Cherrington, N. J., Choudhuri, S., Hartley, D. P., and Klaassen, C. D. (2002). Gender-specific and developmental influences on the expression of rat organic anion transporters. *J. Pharmacol. Exp. Ther.* **301**, 145–151.
- Bus, J. S., and Hammond, L. E. (2007). Regulatory progress, toxicology, and public concerns with 2,4-D: Where do we stand after two decades? *Crop Prot.* **26**, 266–269.
- Carmichael, N. G., Barton, H. A., Boobis, A. R., Cooper, R. L., Dellarco, V. L., Doerrer, N. G., Fenner-Crisp, P. A., Doe, J. E., Lamb, J. C., 4th, and Pastoor, T. P. (2006). Agricultural chemical safety assessment: A multisector approach to the modernization of human safety requirements. *Crit. Rev. Toxicol.* **36**, 1–7.
- Carney, E. W., Zablotny, C. L., Marty, M. S., Crissman, J. W., Anderson, P., Woolhiser, M., and Holsapple, M. (2004). The effects of feed restriction during in utero and postnatal development in rats. *Toxicol. Sci.* **82**, 237–249.
- Cederbrant, K., Marcusson-Ståhl, M., Condevaux, F., and Descotes, J. (2003). NK-cell activity in immunotoxicity drug evaluation. *Toxicology* **185**, 241–250.
- Chapin, R. E., Gulati, D. K., Barnes, L. H., and Teague, J. L. (1993). The effects of feed restriction on reproductive function in Sprague-Dawley rats. *Fundam. Appl. Toxicol.* **20**, 23–29.
- Chapin, R. E., Sloane, R. A., and Haseman, J. K. (1997). The relationships among reproductive endpoints in Swiss mice, using the reproductive assessment by Continuous Breeding database. *Fundam. Appl. Toxicol.* **38**, 129–142.
- Charles, J. M., Cunny, H. C., Wilson, R. D., and Bus, J. S. (1996). Comparative subchronic studies on 2,4-dichlorophenoxyacetic acid, amine, and ester in rats. *Fundam. Appl. Toxicol.* **33**, 161–165.
- Clark, R. L. (1999). Endpoints of reproductive system development. In *An Evaluation and Interpretation of Reproductive Endpoints for Human Risk Assessment*, (Daston, G. and Kimmel, C., eds.) pp. 27–62. International Life Sciences Institute, Health and Environmental Science Institute, Washington, DC.
- Cooper, R. L., Chadwick, R. W., Rehnberg, G. L., Goldman, J. M., Booth, K. C., Hein, J. F., and McElroy, W. K. (1989). Effect of lindane on hormonal control of reproductive function in the female rat. *Toxicol. Appl. Pharmacol.* **99**, 384–394.
- Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerrer, N. G., Eisenbrandt, D. L., Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F., et al. (2006). A tiered approach to life stages testing for agricultural chemical safety assessment. *Crit. Rev. Toxicol.* **36**, 69–98.
- Doe, J. E., Boobis, A. R., Blacker, A., Dellarco, V., Doerrer, N. G., Franklin, C., Goodman, J. L., Kronenberg, J. M., Lewis, R., McConnell, E. E., et al. (2006). A tiered approach to systemic toxicity testing for agricultural chemical safety assessment. *Crit. Rev. Toxicol.* **36**, 37–68.
- Döhler, K. D., Wong, C. C., and von zur Mühlen, A. (1979). The rat as model for the study of drug effects on thyroid function: Consideration of methodological problems. *Pharmacol. Ther. B* **5**, 305–318.
- Duffard, R., Garcia, G., Rosso, S., Bortolozzi, A., Madariaga, M., di Paolo, O., and Evangelista de Duffard, A. M. (1996). Central nervous system myelin deficit in rats exposed to 2,4-dichlorophenoxyacetic acid throughout lactation. *Neurotoxicol. Teratol.* **18**, 691–696.
- Fegert, I., Billington, R., Botham, P., Carney, E., FitzGerald, R. E., Hanley, T., Lewis, R., Marty, M. S., Schneider, S., Sheets, L. P., et al. (2012). Feasibility of the extended one-generation reproductive toxicity study (OECD 443). *Reprod. Toxicol.* **34**, 331–339.
- Florsheim, W. H., and Velcoff, S. M. (1962). Some effects of 2,4-dichlorophenoxyacetic acid on thyroid function in the rat: Effects on iodine accumulation. *Endocrinology* **71**, 1–6.
- Florsheim, W. H., Velcoff, S. M., and Williams, A. D. (1963). Some effects of 2,4-dichlorophenoxyacetic acid on thyroid function in the rat: Effect on peripheral thyroxine. *Endocrinology* **72**, 327–333.
- Frigerio, C., Schutz, Y., Prentice, A., Whitehead, R., and Jéquier, E. (1991). Is human lactation a particularly efficient process? *Eur. J. Clin. Nutr.* **45**, 459–462.
- Gallavan, R. H., Jr., Holson, J. F., Stump, D. G., Knapp, J. F., and Reynolds, V. L. (1999). Interpreting the toxicologic significance of alterations in anogenital distance: Potential for confounding effects of progeny body weights. *Reprod. Toxicol.* **13**, 383–390.
- Garabrant, D. H., and Philbert, M. A. (2002). Review of 2,4-dichlorophenoxyacetic acid (2,4-D) epidemiology and toxicology. *Crit. Rev. Toxicol.* **32**, 233–257.
- Gorzinski, S. J., Kociba, R. J., Campbell, R. A., Smith, F. A., Nolan, R. J., and Eisenbrandt, D. L. (1987). Acute, pharmacokinetic, and subchronic toxicological studies of 2,4-dichlorophenoxyacetic acid. *Fundam. Appl. Toxicol.* **9**, 423–435.
- Gray, L., Ostby, J., Ferrell, J., Sigmon, R., Cooper, R., Linder, R., Rehnberg, G., Goldman, J., and Laskey, J. (1989). Correlation of sperm and endocrine measures with reproductive success in rodents. In *Progress in Clinical and Biological Research* (E. J. Burger Ed.), Vol. 302, pp. 193–209. Liss, New York, NY.
- Grigor, M. R., Allan, J. E., Carrington, J. M., Carne, A., Geursen, A., Young, D., Thompson, M. P., Haynes, E. B., and Coleman, R. A. (1987). Effect of

- dietary protein and food restriction on milk production and composition, maternal tissues and enzymes in lactating rats. *J. Nutr.* **117**, 1247–1258.
- Hanley, T. R., Jr. and Watanabe, P. G. (1985). Measurement of solid feed consumption patterns in neonatal rats by ¹⁴¹Ce-radiolabeled microspheres. *Toxicol. Appl. Pharmacol.* **77**, 496–500.
- Hasegawa, M., Kusuhara, H., Endou, H., and Sugiyama, Y. (2003). Contribution of organic anion transporters to the renal uptake of anionic compounds and nucleoside derivatives in rat. *J. Pharmacol. Exp. Ther.* **305**, 1087–1097.
- Hasegawa, M., Kusuhara, H., Sugiyama, D., Ito, K., Ueda, S., Endou, H., and Sugiyama, Y. (2002). Functional involvement of rat organic anion transporter 3 (rOat3; Slc22a8) in the renal uptake of organic anions. *J. Pharmacol. Exp. Ther.* **300**, 746–753.
- Hays, S. M., Aylward, L. L., Driver, J., Ross, J., and Kirman, C. (2012). 2,4-D exposure and risk assessment: Comparison of external dose and biomonitoring based approaches. *Regul. Toxicol. Pharmacol.* **64**, 481–489.
- Health Canada PACR (2005) Re-evaluation of the lawn and turf uses of (2,4-dichlorophenoxy) acetic acid [2,4-D]. Proposed Acceptability for Continuing Registration. PACR2005-1. 21 February 2005.
- Health Canada PACR (2007) Re-evaluation of the agricultural, forestry, aquatic and industrial site uses of (2,4-dichlorophenoxy) acetic acid [2,4-D]. Proposed Acceptability for Continuing Registration. PACR2007-06. 19 June 2007.
- Hirshfield, A. N. (1987). Histological assessment of follicular development and its applicability to risk assessment. *Reprod. Toxicol.* **1**, 71–79.
- Hotchkiss, A. K., Parks-Saldutti, L. G., Ostby, J. S., Lambright, C., Furr, J., Vandenbergh, J. G., and Gray, L. E., Jr. (2004). A mixture of the “antiandrogens” linuron and butyl benzyl phthalate alters sexual differentiation of the male rat in a cumulative fashion. *Biol. Reprod.* **71**, 1852–1861.
- Jahnke, G. D., Choksi, N. Y., Moore, J. A., and Shelby, M. D. (2004). Thyroid toxicants: Assessing reproductive health effects. *Environ. Health Perspect.* **112**, 363–368.
- Janer, G., Hakkert, B. C., Piersma, A. H., Vermeire, T., and Slob, W. (2007). A retrospective analysis of the added value of the rat two-generation reproductive toxicity study versus the rat subchronic toxicity study. *Reprod. Toxicol.* **24**, 103–113.
- Kim, G. G., Donnenberg, V. S., Donnenberg, A. D., Gooding, W., and Whiteside, T. L. (2007). A novel multiparametric flow cytometry-based cytotoxicity assay simultaneously immunophenotypes effector cells: Comparisons to a 4 h ⁵¹Cr-release assay. *J. Immunol. Methods* **325**, 51–66.
- Korenbrot, C. C., Huhtaniemi, I. T., and Weiner, R. I. (1977). Preputial separation as an external sign of pubertal development in the male rat. *Biol. Reprod.* **17**, 298–303.
- Ladics, G. S., Smith, C., Bunn, T. L., Dietert, R. R., Anderson, P. K., Wescinski, C. M., and Holsapple, M. P. (2000). Characterization of an approach to developmental immunotoxicology assessment in the rat using SRBC as the antigen. *Toxicol. Methods* **10**, 283–311.
- Langhans, B., Ahrendt, M., Nattermann, J., Sauerbruch, T., and Spengler, U. (2005). Comparative study of NK cell-mediated cytotoxicity using radioactive and flow cytometric cytotoxicity assays. *J. Immunol. Methods* **306**, 161–168.
- Larsen, P. R., and Ingbar, S. H. (1992). The thyroid gland. In *Williams Textbook on Endocrinology* (J. D. Wilson and D. W. Foster, Eds.), pp. 357–487. W.B. Saunders Company, Philadelphia, PA.
- Lee, K., Johnson, V. L., and Blakley, B. R. (2001). The effect of exposure to a commercial 2,4-D formulation during gestation on the immune response in CD-1 mice. *Toxicology* **165**, 39–49.
- Loveless, S. E., Ladics, G. S., Smith, C., Holsapple, M. P., Woolhiser, M. R., White, K. L., Jr, Musgrove, D. L., Smialowicz, R. J., and Williams, W. (2007). Interlaboratory study of the primary antibody response to sheep red blood cells in outbred rodents following exposure to cyclophosphamide or dexamethasone. *J. Immunotoxicol.* **4**, 233–238.
- Luke, M. C., and Coffey, D. S. (1994). The male sex accessory tissues. Structure, androgen action, and physiology. In *Physiology of Reproduction* (E. Knobil and J. D. Neill, Eds.), pp. 1435–1487. Raven Press, New York, NY.
- Luster, M. I., Portier, C., Pait, D. G., White, K. L., Jr, Gennings, C., Munson, A. E., and Rosenthal, G. J. (1992). Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. *Fundam. Appl. Toxicol.* **18**, 200–210.
- Mangelsdorf, I., Buschmann, J., and Orthen, B. (2003). Some aspects relating to the evaluation of the effects of chemicals on male fertility. *Regul. Toxicol. Pharmacol.* **37**, 356–369.
- Mann, C. M., Peachee, V. L., Trimmer, G. W., Lee, J. E., Twerdok, L. E., and White, K. L., Jr. (2008). Immunotoxicity evaluation of jet a jet fuel in female rats after 28-day dermal exposure. *J. Toxicol. Environ. Health. A* **71**, 495–504.
- Marcusson-Ståhl, M., and Cederbrant, K. (2003). A flow-cytometric NK-cytotoxicity assay adapted for use in rat repeated dose toxicity studies. *Toxicology* **193**, 269–279.
- Marty, M. S., Allen, B., Chapin, R. E., Cooper, R., Daston, G. P., Flaws, J. A., Foster, P. M., Makris, S. L., Mylchreest, E., Sandler, D., et al. (2009). Interlaboratory control data for reproductive endpoints required in the OPPTS 870.3800/OECD 416 reproduction and fertility test. *Birth Defects Res. B. Dev. Reprod. Toxicol.* **86**, 470–489.
- Marty, M. S., Johnson, K. A., and Carney, E. W. (2003). Effect of feed restriction on Hershberger and pubertal male assay endpoints. *Birth Defects Res. B. Dev. Reprod. Toxicol.* **68**, 363–374.
- Mattsson, J. L., Charles, J. M., Yano, B. L., Cuny, H. C., Wilson, R. D., and Bus, J. S. (1997). Single-dose and chronic dietary neurotoxicity screening studies on 2,4-dichlorophenoxyacetic acid in rats. *Fundam. Appl. Toxicol.* **40**, 111–119.
- Mattsson, J. L., Johnson, K. A., and Albee, R. R. (1986). Lack of neuropathologic consequences of repeated dermal exposure to 2,4-dichlorophenoxyacetic acid in rats. *Fundam. Appl. Toxicol.* **6**, 175–181.
- McClain, R. M. (1995). Mechanistic considerations for the relevance of animal data on thyroid neoplasia to human risk assessment. *Mutat. Res.* **333**, 131–142.
- McIntyre, B. S., Barlow, N. J., and Foster, P. M. (2001). Androgen-mediated development in male rat offspring exposed to flutamide in utero: Permanence and correlation of early postnatal changes in anogenital distance and nipple retention with malformations in androgen-dependent tissues. *Toxicol. Sci.* **62**, 236–249.
- Munro, I. C., Carlo, G. L., Orr, J. C., Sund, K. G., Wilson, R. M., Kennepohl, E., Lynch, B. S., Jablinske, M., and Lee, N. L. (1992). A comprehensive, integrated review and evaluation of the scientific evidence relating to the safety of the herbicide 2,4-D. *J. Amer. Coll. Toxicol.* **11**, 559–664.
- OECD (2011). *Guidance Document on the Current Implementation of Internal Triggers in the Extended One Generation Reproductive Toxicity Study in the United States and Canada*, Series on Testing and Assessment, No. 117. ENV/JM/MONO(2011)21. OECD, Paris, France.
- OECD (2013). *Guidance Document Supporting OECD Test Guideline 443 on the Extended One-Generation Reproductive Toxicity Text*. Series on Testing and Assessment, No. 151. ENV/JM/MONO(2013)10. OECD, Paris, France. Accessed at: [http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=ENV/JM/MONO\(2013\)10&doclanguage=en](http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=ENV/JM/MONO(2013)10&doclanguage=en). Accessed August 19, 2013.
- Pelletier, O., Ritter, L., Caron, J., and Somers, D. (1989). Disposition of 2,4-dichlorophenoxyacetic acid dimethylamine by Fischer 344 rats dosed orally and dermally. *J. Toxicol. Environ. Health* **28**, 221–234.
- Petersen, J. C., Morrissey, R. L., Saunders, D. R., Pavkov, K. L., Luempert, L. G., 3rd, Turnier, J. C., Matheson, D. W., and Schwartz, D. R. (1996). A 2-year comparison study of Crl:CD BR and Hsd:Sprague-Dawley SD rats. *Fundam. Appl. Toxicol.* **33**, 196–211.

- Piersma, A. H., Rorije, E., Beekhuijzen, M. E., Cooper, R., Dix, D. J., Heinrich-Hirsch, B., Martin, M. T., Mendez, E., Muller, A., Paparella, M., *et al.* (2011). Combined retrospective analysis of 498 rat multi-generation reproductive toxicity studies: On the impact of parameters related to F1 mating and F2 offspring. *Reprod. Toxicol.* **31**, 392–401.
- Pizarro, N., Farré, M., Pujadas, M., Peiró, A. M., Roset, P. N., Joglar, J., and de la Torre, R. (2004). Stereochemical analysis of 3,4-methylenedioxymethamphetamine and its main metabolites in human samples including the catechol-type metabolite (3,4-dihydroxymethamphetamine). *Drug Metab. Dispos.* **32**, 1001–1007.
- Prentice, A. M., and Whitehead, R. G. (1987). The energetics of human reproduction. In *Reproductive Energetics in Mammals* (A. Loudon and T. Racey, Eds.), pp. 275–304. Oxford University Press, Oxford, England.
- Regan, K. S., Cline, J. M., Creasy, D., Davis, B., Foley, G. L., Lanning, L., Latendresse, J. R., Makris, S., Morton, D., Rehm, S., *et al.*; STP Ovary Evaluation Working Group. (2005). STP position paper: Ovarian follicular counting in the assessment of rodent reproductive toxicity. *Toxicol. Pathol.* **33**, 409–412.
- Roberts, A., and Renwick, A. G. (1989). The pharmacokinetics and tissue concentrations of cyclohexylamine in rats and mice. *Toxicol. Appl. Pharmacol.* **98**, 230–242.
- Rorije, E., Muller, A., Beekhuijzen, M. E., Hass, U., Heinrich-Hirsch, B., Paparella, M., Schenk, E., Ulbrich, B., Hakkert, B. C., and Piersma, A. H. (2011). On the impact of second generation mating and offspring in multi-generation reproductive toxicity studies on classification and labelling of substances in Europe. *Regul. Toxicol. Pharmacol.* **61**, 251–260.
- Rosso, S. B., García, G. B., Madariaga, M. J., Evangelista de Duffard, A. M., and Duffard, R. O. (2000). 2,4-Dichlorophenoxyacetic acid in developing rats alters behaviour, myelination and regions brain gangliosides pattern. *Neurotoxicology* **21**, 155–163.
- Saghir, S. A., Bartels, M. J., Rick, D. L., McCoy, A. T., Rasoulpour, R. J., Ellis-Hutchings, R. G., Sue Marty, M., Terry, C., Bailey, J. P., Billington, R., *et al.* (2012). Assessment of diurnal systemic dose of agrochemicals in regulatory toxicity testing—an integrated approach without additional animal use. *Regul. Toxicol. Pharmacol.* **63**, 321–332.
- Saghir, S. A., Marty, M. S., Zablony, C. L., Passage, J. K., Perala, A. W., Bus, J. S., Neal, B. H., and Hammond, L. (2013). Life-stage-dependent dietary toxicokinetics of 2,4-dichlorophenoxyacetic acid (2,4-D) in rats: Implications for toxicity test design and interpretation. *Toxicol. Sci.* (accepted).
- Saghir, S. A., Mendrala, A. L., Bartels, M. J., Day, S. J., Hansen, S. C., Sushynski, J. M., and Bus, J. S. (2006). Strategies to assess systemic exposure of chemicals in subchronic/chronic diet and drinking water studies. *Toxicol. Appl. Pharmacol.* **211**, 245–260.
- Sauerhoff, M. W., Braun, W. H., Blau, G. E., and Gehring, P. J. (1977). The fate of 2,4-dichlorophenoxyacetic acid (2,4-D) following oral administration to man. *Toxicology* **8**, 3–11.
- Slikker, W., Jr, Andersen, M. E., Bogdanffy, M. S., Bus, J. S., Cohen, S. D., Conolly, R. B., David, R. M., Doerrer, N. G., Dorman, D. C., Gaylor, D. W., *et al.* (2004). Dose-dependent transitions in mechanisms of toxicity. *Toxicol. Appl. Pharmacol.* **201**, 203–225.
- Smart, J. L., Massey, R. F., Nash, S. C., and Tonkiss, J. (1987). Effects of early-life undernutrition in artificially reared rats: Subsequent body and organ growth. *Br. J. Nutr.* **58**, 245–255.
- Sweatman, T. W., and Renwick, A. G. (1980). The tissue distribution and pharmacokinetics of saccharin in the rat. *Toxicol. Appl. Pharmacol.* **55**, 18–31.
- Thomas, K. W., Dosemeci, M., Hoppin, J. A., Sheldon, L. S., Croghan, C. W., Gordon, S. M., Jones, M. L., Reynolds, S. J., Raymer, J. H., Akland, G. G., *et al.* (2010). Urinary biomarker, dermal, and air measurement results for 2,4-D and chlorpyrifos farm applicators in the Agricultural Health Study. *J. Expo. Sci. Environ. Epidemiol.* **20**, 119–134.
- Timchalk, C. (2004). Comparative inter-species pharmacokinetics of phenoxy-acetic acid herbicides and related organic acids. Evidence that the dog is not a relevant species for evaluation of human health risk. *Toxicology* **200**, 1–19.
- Tojo, A., Sekine, T., Nakajima, N., Hosoyamada, M., Kanai, Y., Kimura, K., and Endou, H. (1999). Immunohistochemical localization of multispecific renal organic anion transporter 1 in rat kidney. *J. Am. Soc. Nephrol.* **10**, 464–471.
- Ulbrich, B., and Palmer, A. K. (1995). Detection of effects on male reproduction—A literature survey. *J. Am. Coll. Toxicol.* **14**, 293–327.
- U.S. EPA (2013). Supporting Statement for an Information Collection Request (ICR) entitled, “Tier 2 Data Collection for Certain Chemicals under the Endocrine Disruptor Screening Program (EDSP)”. EPA ICR No. 2479.01; Doct ID No. EPA-HQ-OPPT-2013-0171. Washington, DC. June 18, 2013. Available at: <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2013-0171-0002>. Accessed August 6, 2013.
- U.S. EPA RED (2005). *Reregistration Eligibility Decision for 2,4-D*. Office of Pesticide Programs, Washington, DC. EPA 738-R-05-002. June 2005.
- Van den Berg, K. J., van Raaij, J. A., Bragt, P. C., and Notten, W. R. (1991). Interactions of halogenated industrial chemicals with transthyretin and effects on thyroid hormone levels in vivo. *Arch. Toxicol.* **65**, 15–19.
- van Ravenzwaay, B., Hardwick, T. D., Needham, D., Pethen, S., and Lappin, G. J. (2003). Comparative metabolism of 2,4-dichlorophenoxyacetic acid (2,4-D) in rat and dog. *Xenobiotica* **33**, 805–821.
- White, K. L., Jr, Germolec, D. R., Booker, C. D., Hernandez, D. M., McCay, J. A., Delclos, K. B., Newbold, R. R., Weis, C., and Guo, T. L. (2005). Dietary methoxychlor exposure modulates splenic natural killer cell activity, antibody-forming cell response and phenotypic marker expression in F0 and F1 generations of Sprague Dawley rats. *Toxicology* **207**, 271–281.
- Wolf, C. J., Hotchkiss, A., Ostby, J. S., LeBlanc, G. A., and Gray, L. E., Jr. (2002). Effects of prenatal testosterone propionate on the sexual development of male and female rats: A dose-response study. *Toxicol. Sci.* **65**, 71–86.
- Yoon, M., and Barton, H. A. (2008). Predicting maternal rat and pup exposures: How different are they? *Toxicol. Sci.* **102**, 15–32.

Supplemental Data:

AN F₁-EXTENDED ONE GENERATION REPRODUCTIVE TOXICITY STUDY
(EOGRTS) IN CRL:CD(SD) RATS WITH 2,4-DICHLOROPHENOXYACETIC
ACID (2,4-D)

Marty, M.S., Neal, B.H., Zablony, C.L., Yano, B.L., Andrus, A.K., Woolhiser, M.R.,
Boverhof, D.R., Saghir, S.A., Perala, A.W., Passage, J.K., Lawson, M.A., Bus, J.S.,
Lamb, J.C., IV, and L. Hammond

SUPPLEMENTARY DATA DESCRIPTION

Supplemental data includes the following information/data:

Figures S1-S5: Diagrams showing study designs for the various cohorts included in the 2,4-D EOGRTS. Study designs outline timelines and endpoints for the P1 satellite females (Figure S1), F1 Set 1A systemic toxicity and Set 1B neurotoxicity groups (Figure S2), F1 Sets 2A and 2B immunotoxicity groups (Figure S3), Set 3 reproductive toxicity group (Figure S4), and PND 22 weanlings – unselected pups and neuropathology group (Figure S5).

Figures S6-S12: Figures showing the results of neurobehavioral testing and brain morphometry in the F1 Set 1B animals evaluated from PND 54-60, including motor activity (Figure S6), motor activity by epoch (Figure S7), rectal temperature (Figure S8), forelimb grip performance (Figure S9), landing footsplay (Figure S10), acoustic startle response (Figure S11), and brain morphometry (Figure S12).

Figures S13-S14: Figures showing the results of immunotoxicity assessments in the F1 Set 2 animals, including T-cell dependent antibody responses (Set 2A) in Figure S13 and Natural Killer (NK) cell activity (Set 2B) in Figure S14.

Table S1 lists the specific study endpoints that were examined across different life stages in the 2,-D EOGRTS.

Table S2 lists the clinical chemistry and hematology parameters measured in the 2,4-D EOGRTS and the equipment used for these measurements.

Table S3 lists the statistical analyses used for the various data sets collected during the 2,4-D EOGRTS study.

Table S4 provides mathematical formulas used to calculate reproductive indices and colony forming colony (AFC) responses in the sheep red blood cell assay (SRBC).

Table S5 provides data on test material intake across life stages in two cohorts of F1 animals, the set 1A animals exposed to PND 70 and the Set 3 animals exposed to PND 139.

Table S6 shows clinical chemistry data for P1 and F1 Set 1A males and females to evaluate the impact of kidney toxicity.

Table S7 shows results for the reproductive indices and estrous cycle length in the P1 animals.

Tables S8-S11 show a comparison of the endpoints collected in standard studies compared with the 2,4-D EOGRTS and the newly adopted OECD 443 EOGRTS test guideline. Table S8 shows a comparison of the reproductive toxicity test guideline with

reproductive endpoints assessed in the 2,4-D EOGRTS and guideline EOGRTS. Table S9 shows a comparison of the developmental neurotoxicity test guideline with the developmental neurotoxicity endpoints assessed in the 2,4-D EOGRTS and guideline EOGRTS. Table S10 shows a comparison of the immunotoxicity test guideline with the developmental immunotoxicity endpoints assessed in the 2,4-D EOGRTS and guideline EOGRTS. Table S11 shows a comparison of thyroid endpoints in the comparative thyroid guidance document with the thyroid endpoints assessed in the 2,4-D EOGRTS and guideline EOGRTS.

Figure S1. Study Design for P1 Satellite Females in the 2,4-D EOGRTS

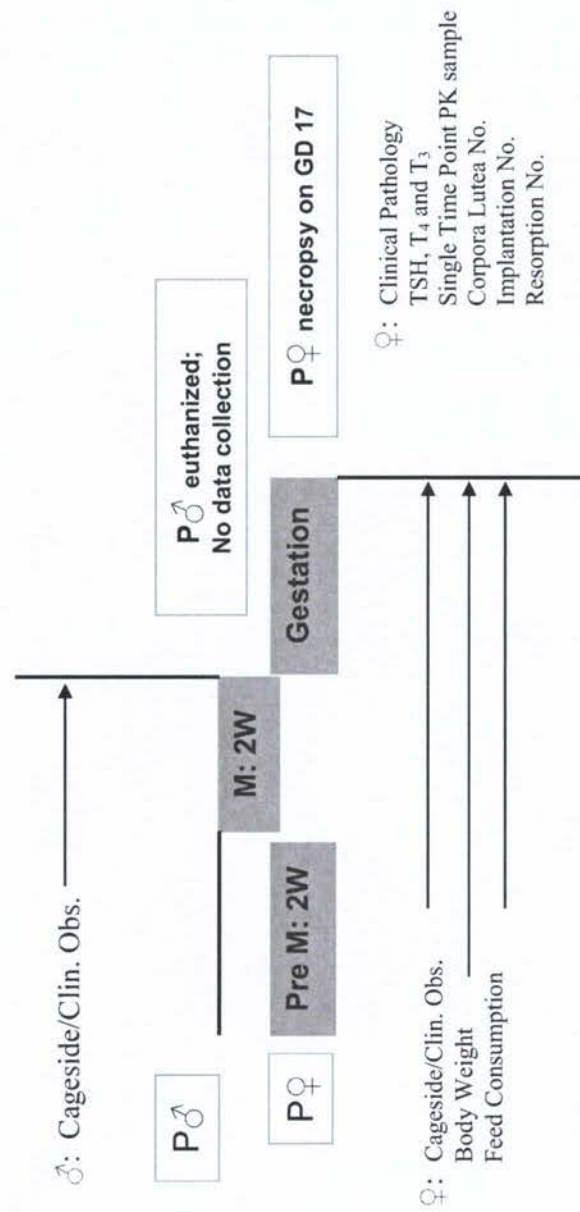
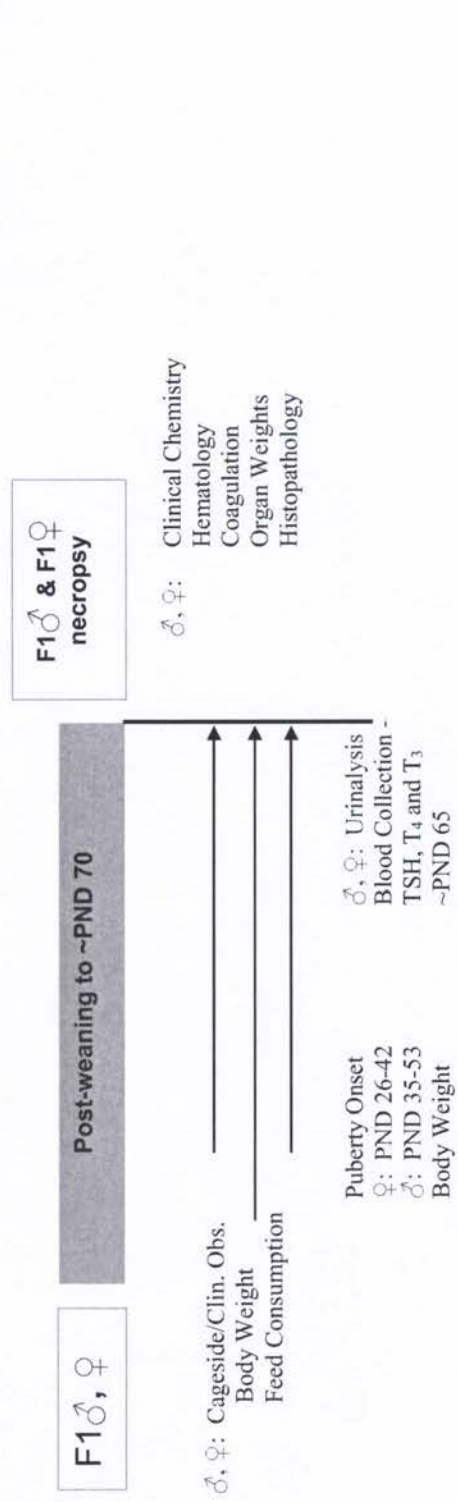


FIGURE S2. Study Design for F1 Generation Set 1a (Systemic Toxicity Group) and 1b (Neurotoxicity Group) in the 2,4-D EOGRTS

Set 1a:



Set 1b:

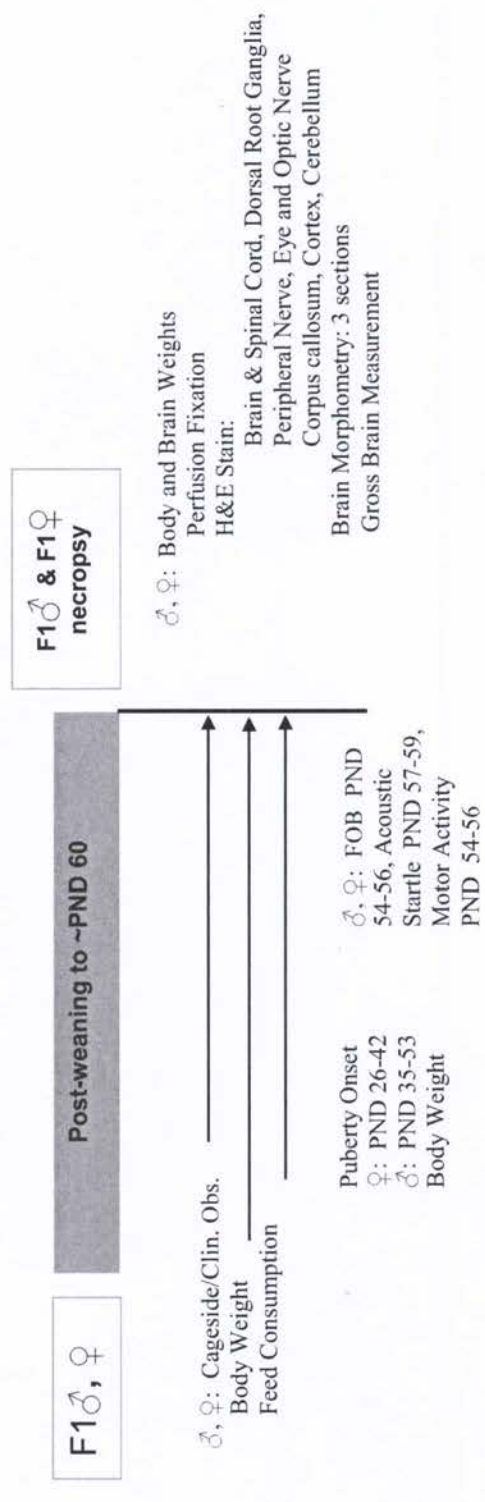
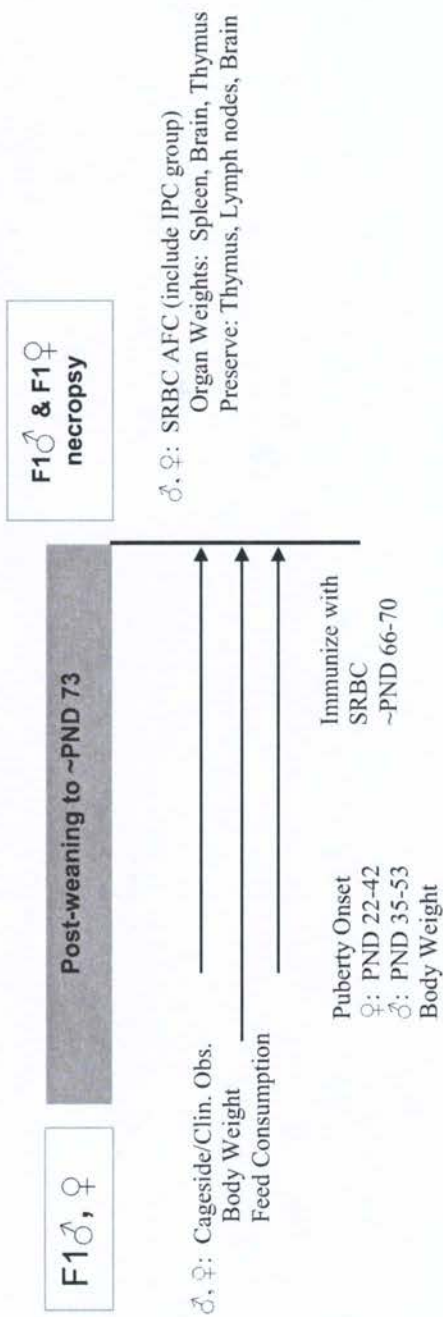


FIGURE S3. Study Design for Set 2 F1 Generation (Immunotoxicity Group) in the 2,4-D EOGRTS

Set 2a:



Set 2b:

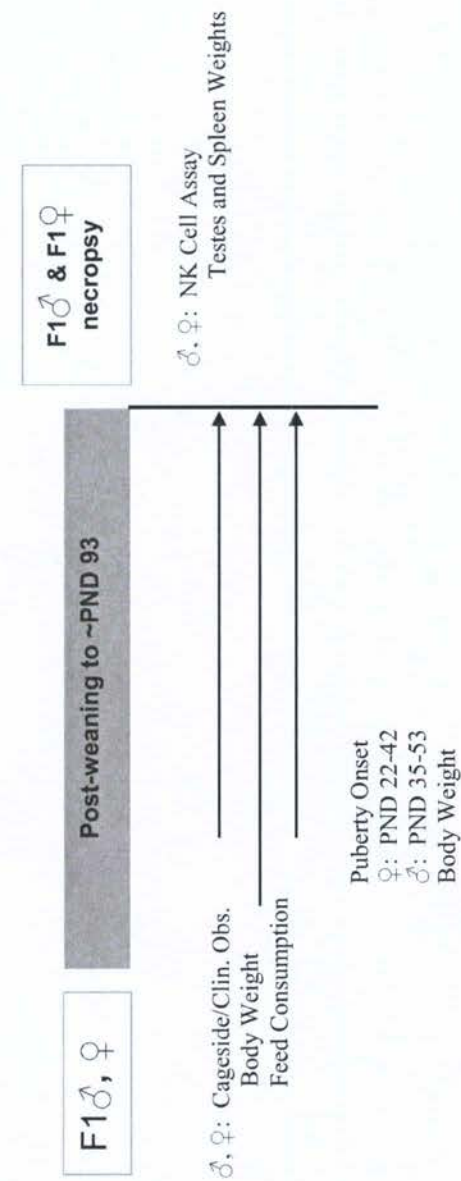


FIGURE S4. Study Design for F1 Generation Set 3 (Reproductive Toxicity Group) in the 2,4-D EOGRTS

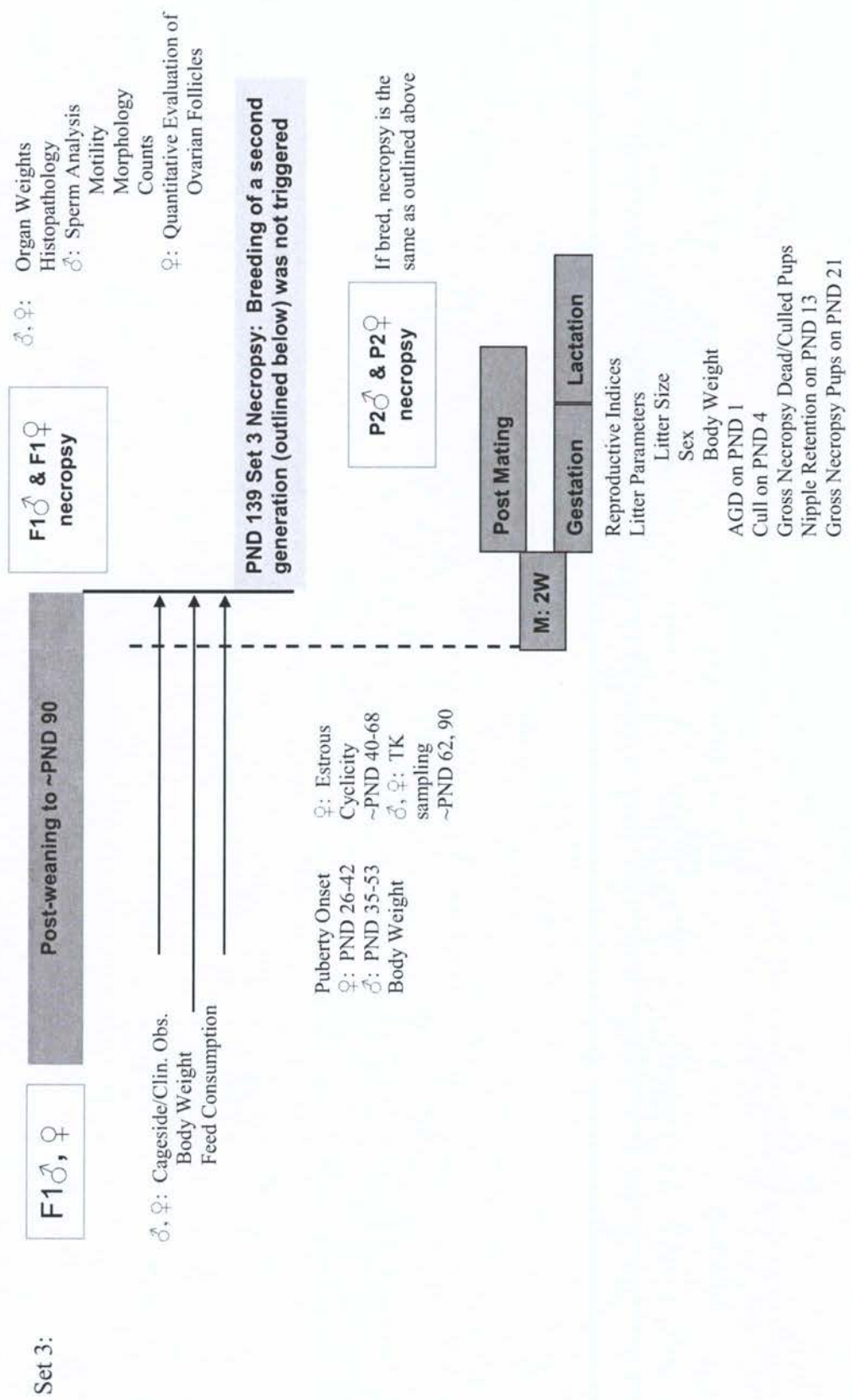
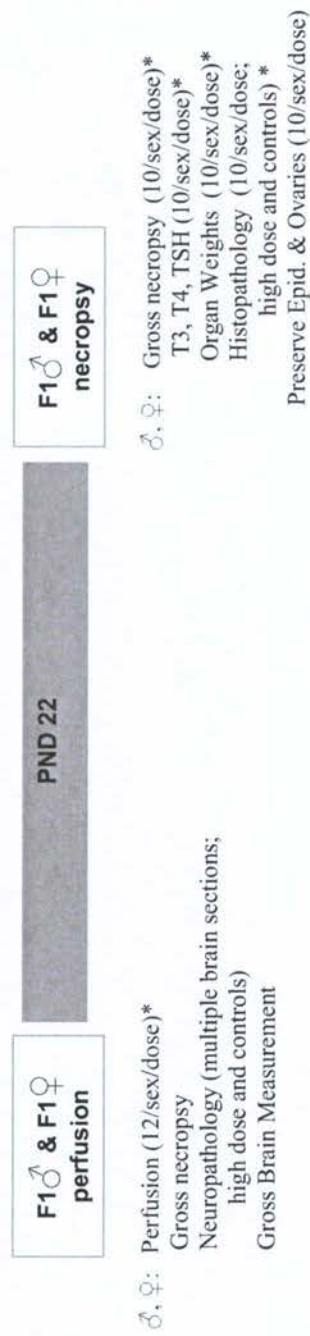


Figure S5. Study Design for PND 22 Neuropathology Group (with Perfusions) and Unselected Weanlings in the 2,4-D EOGRTS



*Using only 1 pup/sex/litter

Figure S6. Total Motor Activity Counts in Set 1B Animals

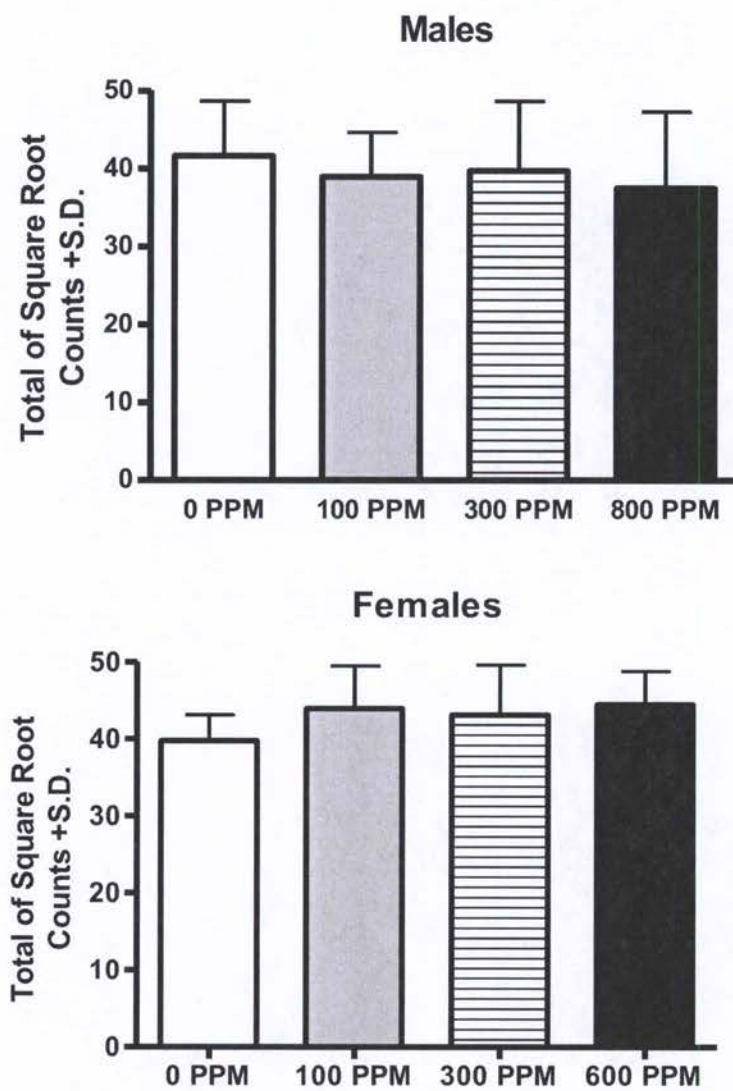


Figure S7. Motor Activity Counts by Epoch in Set 1B Animals

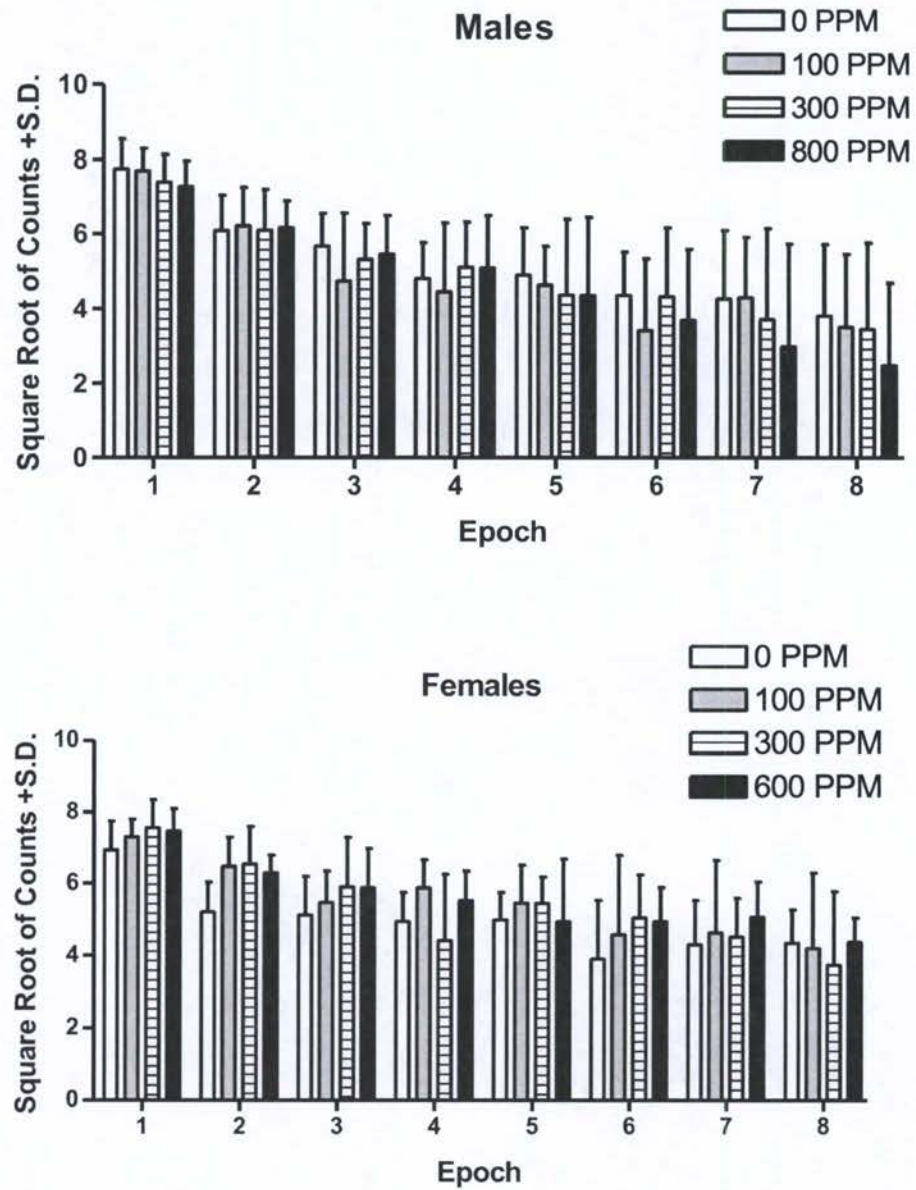


Figure S8. Rectal Temperature in Set 1B Animals

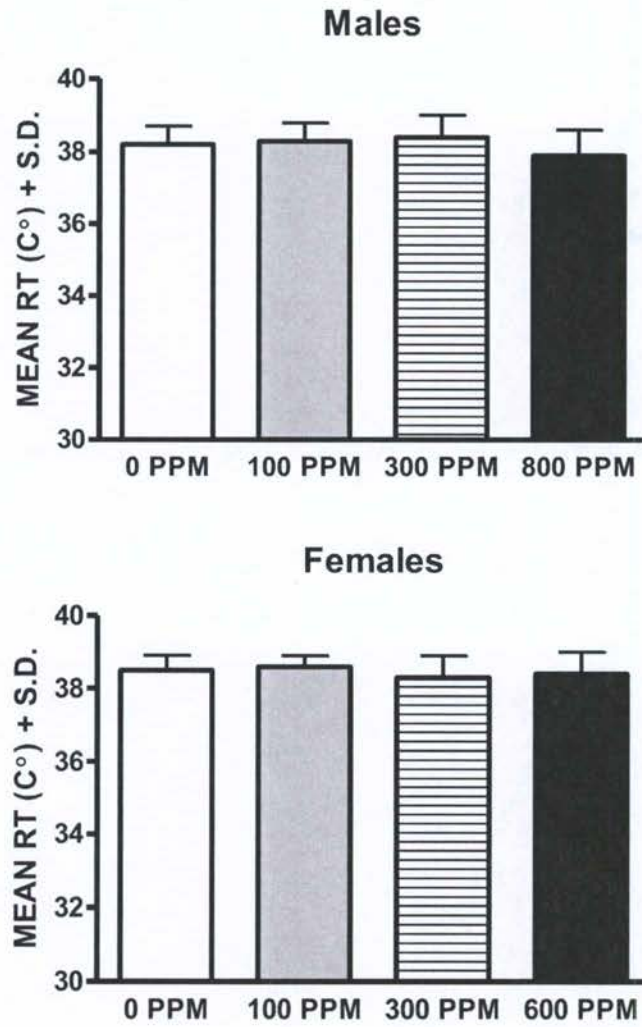


Figure S9. Forelimb Grip Performance in Set 1B Animals

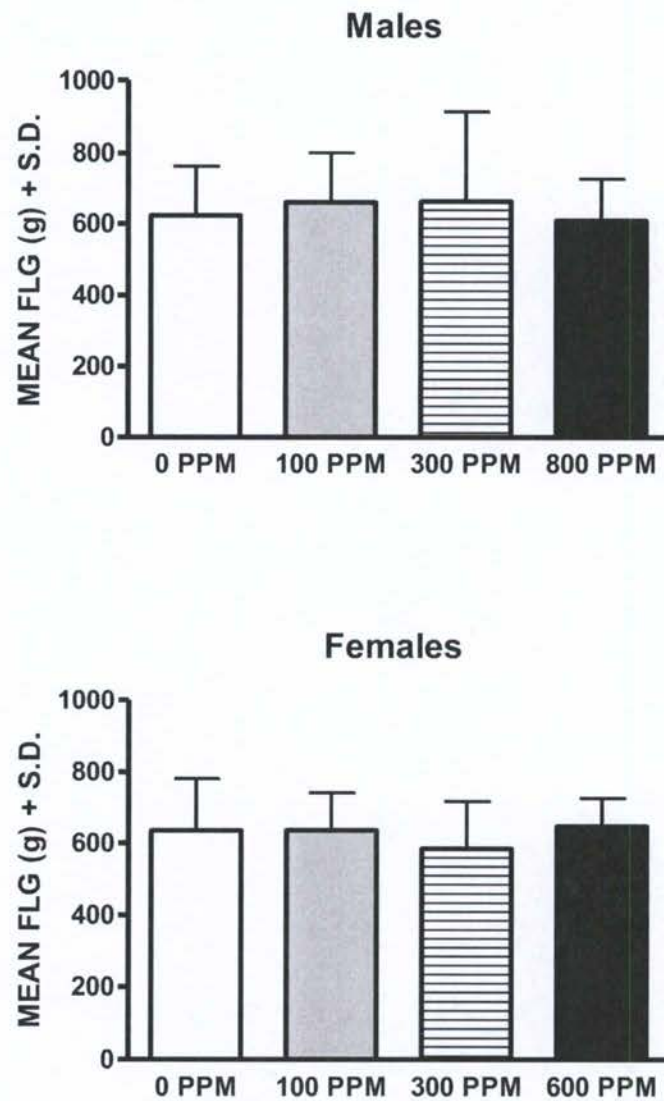


Figure S10. Landing Foot Splay in Set 1B Animals

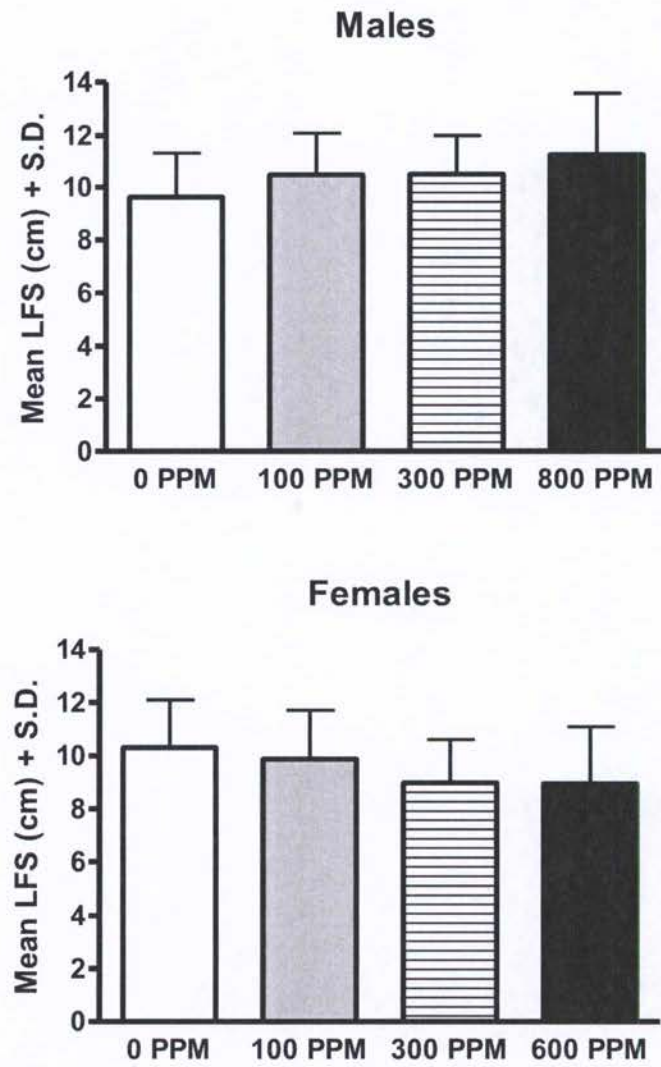


Figure S11. Acoustic Startle Response in Set 1B Animals

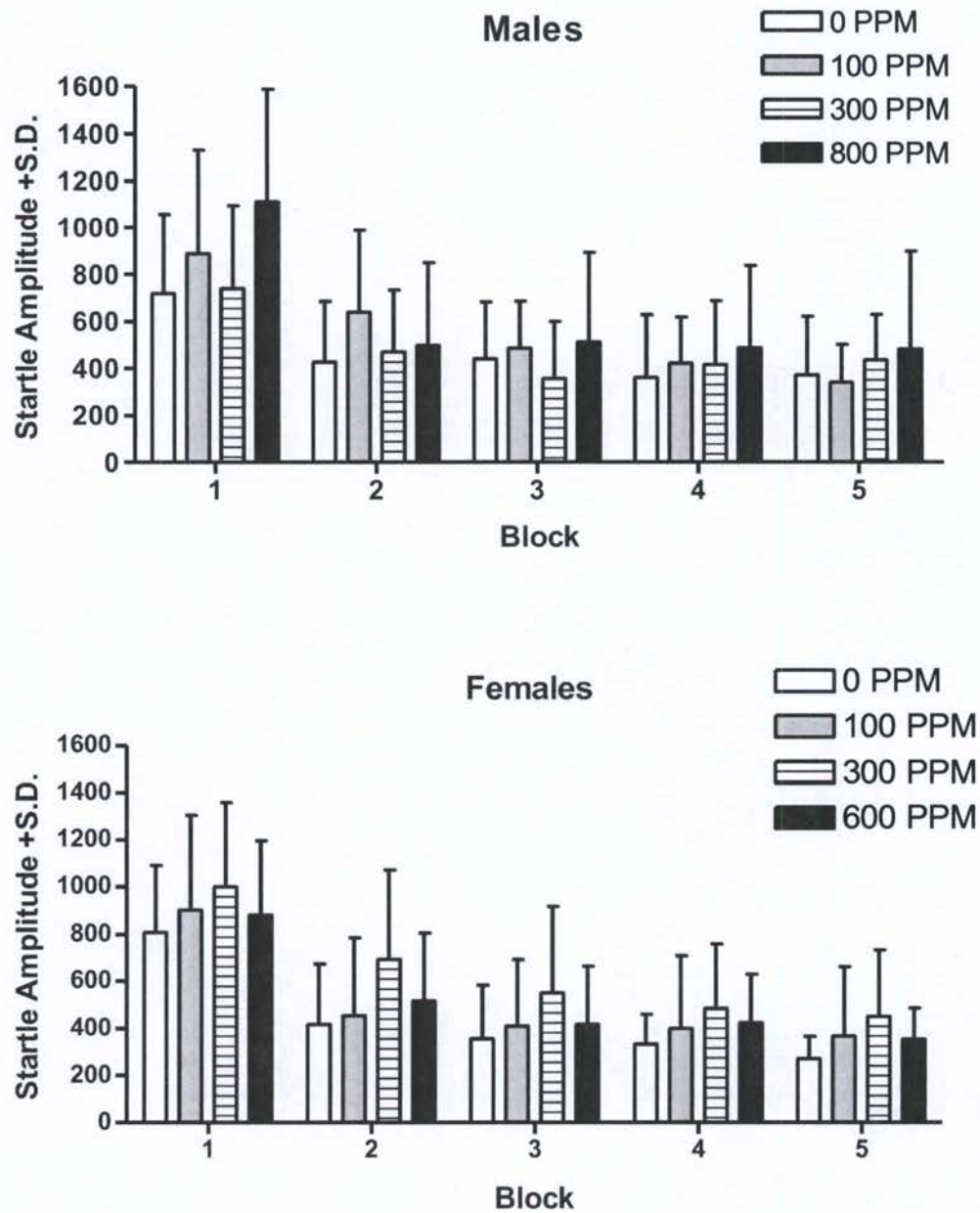


Figure S12. Brain Morphometry Measurements in Set 1B Animals

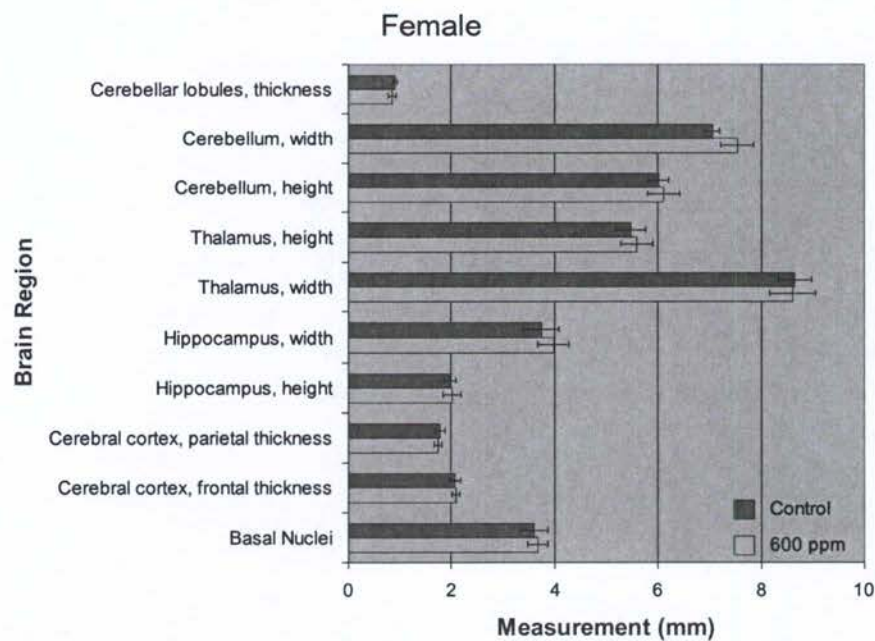
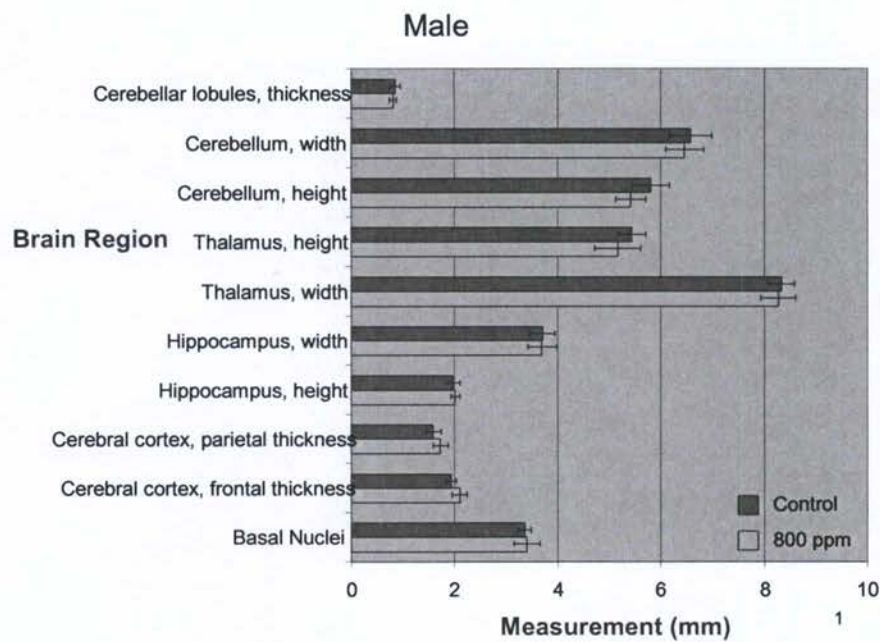


Figure S13. T-Cell Dependent Antibody Response in Set 2A Animals

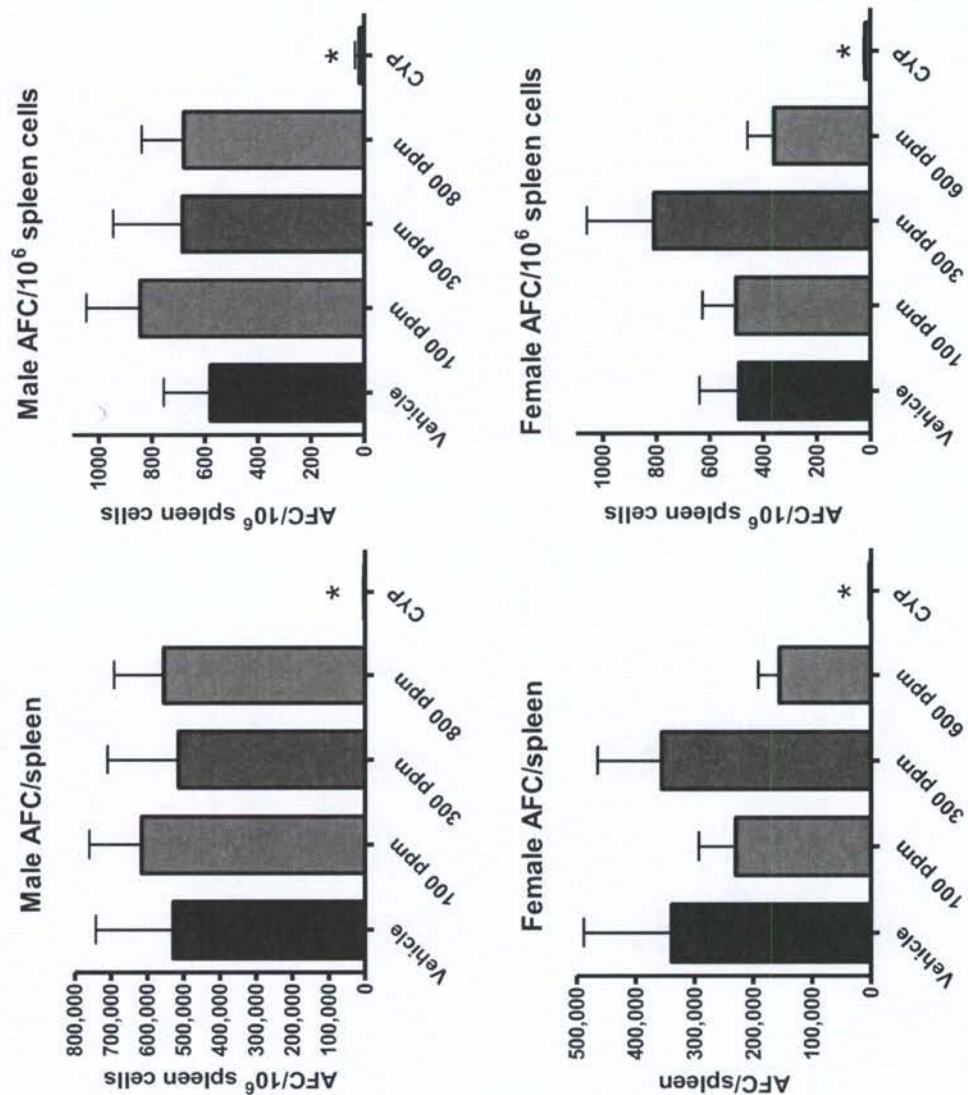


Figure S14. Natural Killer (NK) Cell Activity in Set 2B Animals

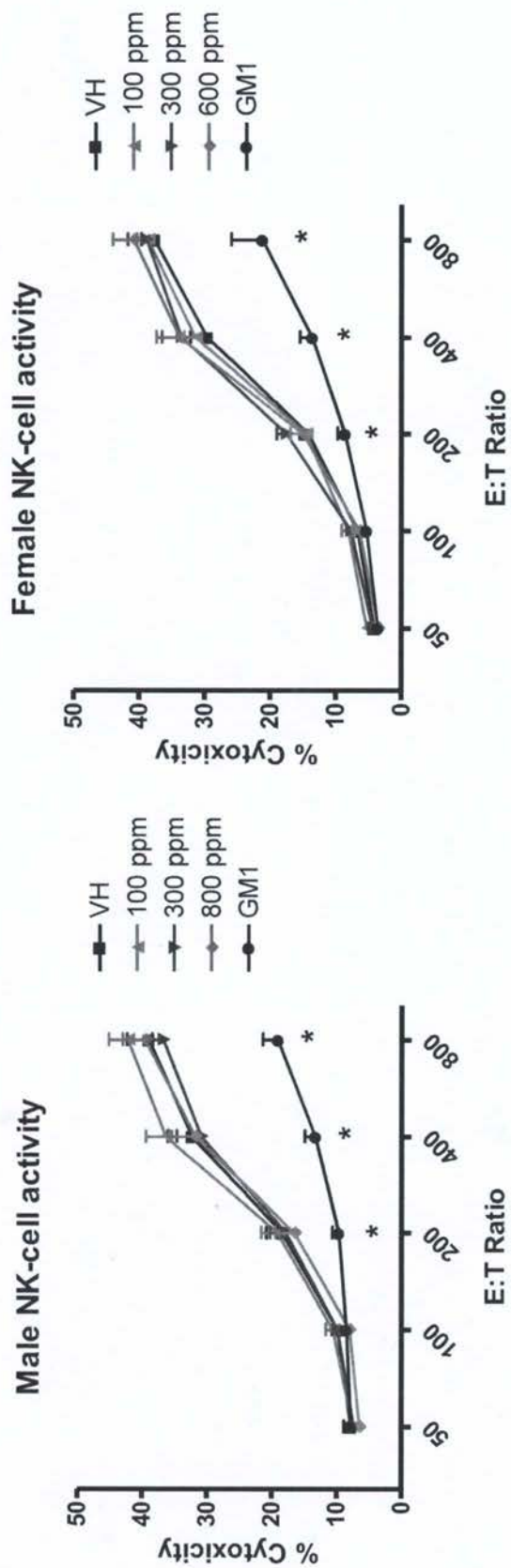


Table S1. Primary Study Endpoints Examined Across Different Life Stages

	P1 Males	P1 Females	P1 Satellite Females	F1 Cull to 10: ~5/sex per litter	Unselected F1 Weanlings (non-perfused)	Unselected F1 Weanlings (perfused)	F1 Set 1a	F1 Set 1b	F1 Set 2a	F1 Set 2b	F1 Set 3
Sampling Plan	>20/dose ¹	>20/dose ¹	12/dose	Cull to 10: ~5/sex per litter	1 pup/sex /litter	1 pup/sex/litter	1 pup/litter	1 pup/litter	1 pup/sex/litter	1 pup/sex/litter	1 pup/sex/litter
N					10/sex/dose	12/sex/dose	10/sex/dose	10/sex/dose	10/sex/dose	10/sex/dose	>20/sex/dose ²
Age at necropsy	~21 weeks ³	LD 22	GD 17	PND 4	PND 22	PND 22	PND 70	PND 60	PND 67-73	PND 87-93	PND 139 ± 1
Blood Collected											
TK											
Thyroid hormones ⁴				X	X						
Clinical chemistry/hematology	X	X	X				X				
Urinalysis	X						X				
Necropsy with gross exam			X		X						
Terminal Body Weights	X	X	X	X	X	X	X	X	X	X	X
Organ Weights											
Adrenal	X	X			X		X				X
Brain	X	X			X		X	X	X		X
Epididymides	X						X				X
Heart							X				
Kidney	X	X			X		X				X
Liver	X	X			X		X				X
Ovaries		X					X				X
Pituitary ²	X	X			X		X				X

¹ 10-12/sex examined histopathologically from control and high-dose groups; tissues with exposure-related changes were examined in the low- and middle-dose groups as well as reproductive organs from animals that failed to mate or were infertile.

² 10/sex/dose examined histopathologically except for testes and epididymides, which were examined in all Set 3 control and high-dose males; tissues with exposure-related changes were examined in the low- and middle-dose groups as well as reproductive organs from animals that failed to mate or were infertile.

³ Necropsied after ~11 weeks of exposure

⁴ T₃, T₄ and TSH

Table S2. Equipment Used and Clinical Chemistry/Hematology Parameters Measured

Equipment	Parameter(s)
Bayer Advia 120 Hematology Analyzer (Siemens Medical Solutions Diagnostics, Tarrytown, New York)	Hematocrit, Hemoglobin Concentration, Red Blood Cell Count, Total White Blood Cell Count, Differential White Blood Cell Count, Platelet Count, Reticulocyte Count, Mean Corpuscular hemoglobin, Mean Corpuscular volume, Mean Corpuscular Hemoglobin Concentration
Instrumentation Laboratory ACL9000 (Beckman Coulter, Inc., Miami, Florida)	Prothrombin Time
Hitachi 912 Clinical Chemistry Analyzer (Roche Diagnostics, Indianapolis, Indiana)	Albumin, Cholesterol, Creatinine, Electrolytes (Sodium, Potassium, Phosphorus, Chloride, and Calcium), Globulin, Glucose, Total Bilirubin, Total Protein, Triglycerides, Urea Nitrogen
Multistix® Reagent Strips and Bayer HealthCare Clinitek Advantus Analyzer (Siemens Medical Solutions Diagnostics, Tarrytown, New York)	Urine pH, Bilirubin, Glucose, Protein, Ketones, Blood, Urobilinogen
Radioimmunoassays ¹ Anilytics Incorporated (Gaithersburg, Maryland)	T ₄ , T ₃ and TSH

¹ Radioimmunoassays used standards from the National Hormone and Pituitary Program (University of California – Los Angeles, CA).

Table S3. Statistical Analyses

Parameter	Statistical Analysis
Body weights, body weight gains, litter mean body weights, feed consumption, litter mean age and body weight at vaginal opening or preputial separation, anogenital distance (absolute and relative to cubed root of body weight), follicle count, sperm count, percent total and progressively motile sperm, mean estrous cycle length, T ₄ , T ₃ and TSH levels, urine volume, urine specific gravity, clinical chemistry data (excluding globulin and albumin:globulin ratio), coagulation and hematology data (excluding differential WBC counts and RBC indices) organ weights (absolute and relative), spleen cell counts, SRBC AFC data and cytotoxicity data (if no effect was found for SRBC assay)	Bartlett's test for equality of variance ($\alpha = 0.01$, Winer, 1971), parametric analysis of variance (ANOVA; $\alpha = 0.05$, Steel and Torrie, 1960); If significant, Dunnett's test ($\alpha = 0.05$, Winer, 1971)
Body weights, body weight gains, anogenital distance, sperm count, percent total and progressively motile sperm, percent abnormal sperm, number of days per estrous cycle, organ weights, and SRBC AFC data	Bartlett's test for equality of variance, nonparametric analysis of variance ($\alpha = 0.05$, Hollander and Wolfe, 1973); If significant, Wilcoxon Rank-Sum test ($\alpha = 0.05$, Hollander and Wolfe, 1973) with Bonferroni's correction (Miller, 1966)
Litter size, corpora lutea number (satellite P1 females), gestation length, average time to mating	Nonparametric analysis of variance ($\alpha = 0.05$); If significant, Wilcoxon Rank-Sum test with Bonferroni's correction
Pup survival indices	Wilcoxon test as modified by Haseman and Hoel (1974); ($\alpha = 0.05$)
Sperm morphology	Arcsine transformed and analyzed using a parametric ANOVA ¹⁷ ($\alpha = 0.05$, Steel and Torrie, 1960); If significant, Dunnett's test ($\alpha = 0.05$, Winer, 1971)
Mating, conception, fertility and gestation indices	Fisher exact probability test ($\alpha = 0.05$; Siegel, 1956) with Bonferroni's correction
Neonatal sex ratio (PND 1)	Binomial distribution test ($\alpha = 0.05$; Steel and Torrie, 1960)
Survival indices, post-implantation loss, and other incidence data among neonates	(Litter as the experimental unit) Censored Wilcoxon test ($\alpha = 0.05$; Hollander and Wolfe, 1973) as modified by Haseman and Hoel (1974) with Bonferroni's correction
Nipple retention	Fisher exact probability test
DCO incidence data (scored observations)	(Data collected at different time points analyzed separately) z-test of proportions ($\alpha = 0.05$; Bruning and Kintz, 1987)

¹⁷ Slides containing less than 200 sperm were excluded from analysis.

Ranked scores for FOB observations	Converted to average scores for each observation (descriptive only)
Incidence tables of FOB observations	z-test of proportions ($\alpha = 0.05$; Bruning and Kintz, 1987) with the factors of sex, severity level and dose level comparisons
FOB body weights, grip performance, rectal temperature, landing foot splay	(analyzed separately by sex; for grip performance and landing foot splay, the average of 3 trials was used); 1-way ANOVA ($\alpha = 0.05$; Steel and Torrie, 1960); If significant, Dunnett's test ($\alpha = 0.05$, Winer, 1971)
MA counts	(analyzed separately by sex) Square root transformed (Pryor et al., 1983); Factorial repeated-measure design, the multivariate approach, with factors of dose and the repeated factor of epoch; Primary interaction was dose-by-epoch and the dose main effect; If either significant, linear contrasts ($\alpha = 0.02$) with the Pillai's Trace p-value used
ASR habituation	(analyzed separately by sex) Peak values for 50 trial sessions averaged into 5 10-trial blocks for each animal; Factorial repeated-measure design, the multivariate approach, with factors of dose and the repeated factor of block; Primary interaction was dose-by-block and the dose main effect; If either significant, linear contrasts ($\alpha = 0.02$) with the Pillai's Trace p-value used
Gross brain measurements and brain morphometry	(analyzed separately by sex) Brain weight and final body weights analyzed by ANOVA; One-way multivariate analysis of variance (MANOVA) with dose as a factor was conducted on 4 gross brain measurements (cerebrum and cerebellum length and width) ($\alpha = 0.05$); If significant, linear contrasts ($\alpha = 0.02$); ANOVA with modified Benjamini- Hochberg method for each microscopic brain measurement for localized effects (Ellis et al., 2000)
TK data	Descriptive statistics (mean \pm standard deviation)

Table S4. Calculation of Reproductive and Immunotoxicity Parameters

Reproductive indices were calculated as follows:

- Female mating index = (No. females with evidence of mating/No. paired) x 100
- Male mating index = (No. males with evidence of mating/No. paired) x 100
- Female conception index = (No. females with evidence of pregnancy/No. mated) x 100
- Male conception index = (No. males siring a litter/No. mated) x 100
- Female fertility index = (No. females with evidence of pregnancy/No. paired) x 100
- Male fertility index = (No. males siring a litter/No. paired) x 100
- Gestation index = (No. females delivering a viable litter/No. females with evidence of pregnancy) x 100
- Gestation survival index = percentage of delivered pups alive at birth
- Post-implantation loss = (No. implants – No. viable offspring)/(No. implants) x 100
- Day 1 or 4 pup survival index = (No. viable pups on day 1 or 4/No. born live) x 100
- Day 7, 14, or 21 pup survival index = (No. viable pups on day 7, 14 or 21/No. after culling) x 100

The SRBC AFC Assay data were calculated as follows:

AFC/ml (original suspension)	=	Number of plaques x 10 x dilution factor
AFC/rats (spleen)	=	AFC/ml x ml of original spleen cell suspension
AFC/10 ⁶ spleen cells	=	$\frac{\text{AFC/rat} \times 10^6}{\text{Total spleen cell number}}$

Table S5. Test Material Intake Across Life Stages

Dose (ppm)	Age (days)														
	Set 3 Male Test Material Intake (mg 2,4-D/kg body weight)														
	28-35 ^a	35-42	42-49	49-56	56-63	63-70	70-77	77-84	84-91	91-98	98-105	105-112	112-119	119-126	126-133
100	7.0	12.7	10.7	9.3	8.2	7.1	6.6	6.0	5.6	5.4	5.1	5.0	4.8	4.6	4.5
300	21.4	38.4	32.2	27.9	24.7	21.5	19.5	18.8	17.4	16.6	15.9	15.4	14.8	14.3	14.1
800	57.4	102	86.8	75.6	64.1	58.2	53.3	50.4	45.3	43.1	41.4	40.5	39.5	38.5	37.5
Dose (ppm)	Set 3 Female Test Material Intake (mg 2,4-D/kg body weight)														
	28-35 ^a	35-42	42-49	49-56	56-63	63-70	70-77	77-84	84-91	91-98	98-105	105-112	112-119	119-126	126-133
100	6.9	12.2	10.4	9.6	8.7	8.0	7.6	7.3	6.8	6.6	6.2	6.1	6.0	5.8	5.7
300	21.8	38.3	32.1	28.5	26.5	24.3	23.5	22.3	20.8	20.0	18.9	18.7	18.3	17.8	17.2
600	43.3	76.4	62.8	58.1	52.8	48.1	46.1	45.5	42.1	40.3	38.4	38.1	37.0	35.4	35.9
Dose (ppm)	Set 1A Male Test Material Intake (mg 2,4-D/kg body weight)														
	28-35 ^a	35-42	42-49	49-56	56-63	63-69									
100	7.1	12.8	10.9	9.2	7.7	6.4									
300	21.5	39.5	33.6	29.1	24.0	19.0									
800	59.3	106	89.5	77.3	63.6	53.3									
Dose (ppm)	Set 1A Female Test Material Intake (mg 2,4-D/kg body weight)														
	28-35 ^a	35-42	42-49	49-56	56-63	63-69									
100	7.1	12.6	10.8	9.8	8.6	7.1									
300	21.3	38.2	32.5	29.6	25.5	21.3									
600	43.0	77.8	66.1	58.8	49.4	44.1									

^aTest animals received ½ the adult female dietary concentration of 2,4-D during the PND 28-35 interval.

Shaded cells indicate the period when test animals received greater 2,4-D exposures than adults during pre-mating or gestation.

Table S6. Select Clinical Chemistry and Urinalysis Parameters in P1 and Set 1A Animals

Dose Level ppm	Group	UN mg/dL	ALT μ /L	ALP μ /L	AST μ /L	TP g/dL	ALB g/dL	Glob g/dL	A/G ratio	Chol mg/dL	Trig mg/dL	TBil mg/dL	Crea mg/dL	GGT μ /L	Urine specific gravity	Bilirubin (No. pos. and grade)
0	P1 Female	21	65	102	117	6.9	4.0	2.9&	1.4&	72	96	0.1	0.4	<3	NA	NA
100	P1 Female	21	64	96	114	7.0	4.0	3.0&	1.3&	80	89	0.1	0.4	<3	NA	NA
300	P1 Female	21	67	86	118	6.7	3.9	2.9&	1.4&	81	81	0.1	0.4	<3	NA	NA
600	P1 Female	20	79	89	124	7.0	4.0	2.9&	1.4&	79	95	0.1	0.4	<3	NA	NA
0	P1 Male	16	41	89	95	7.0	3.8	3.3&	1.2&	58	52	0.1	0.4	<3	1.044	1/10 (+)
100	P1 Male	16	38	78	92	6.9	3.7	3.2&	1.2&	56	47	0.1	0.4	<3	1.046	1/10 (++)
300	P1 Male	17	42	74	94	6.9	3.7	3.1&	1.2&	54	42	0.1	0.4	<3	1.039	1/10 (+)
800	P1 Male	17	74	74	120	6.6*	3.6*	3.0&	1.2&	59	47	0.1	0.4	<3	1.041	3/10 (+)
0	Set 1A Female	18	32	117	93	6.8	4.0	2.7&	1.5&	57	40	0.1	0.3	<3	1.039	1/10 (+)
100	Set 1A Female	18	34	125	94	7.0	4.2	2.9&	1.5&	63	34	0.1	0.4	<3	1.042	1/10 (+)
300	Set 1A Female	17	36	115	91	6.8	4.1	2.8&	1.5&	71	41	0.1	0.4	<3	1.032	0/10 (+)
600	Set 1A Female	18	40	119	94	6.8	4.0	2.8&	1.5&	68	57	0.1	0.4	<3	1.031	0/10 (+)
0	Set 1A Male	15	40	156	96	6.4	3.7	2.7&	1.4&	50	39	0.1	0.3	<3	1.048	0/10 (+)
100	Set 1A Male	15	37	181	104	6.7	3.8	2.9&	1.4&	54	41	0.1	0.3	<3	1.043	1/10 (+)
300	Set 1A Male	15	42	170	108	6.3	3.6	2.7&	1.4&	51	48	0.1	0.3	<3	1.042	1/10 (+)
800	Set 1A Male	15	47*	186	108	6.4	3.8	2.7&	1.4&	51	51	0.1	0.3	<3	1.044	2/10 (+)

N = Number of replicates = 10 for all parameters except for the 100 ppm Set 1A Female dose group, where n = 9 due to lost specimen during storage.
NA = not applicable; urine samples not collected in P1 females due to lactation.

& - Indicates no statistical comparison of means.

There were no statistical differences from control at $\alpha=0.05$.

* - Statistically different from Control mean by Dunnett's Test, $\alpha = 0.05$.

UN = Urea Nitrogen, ALT = Alanine Aminotransferase, ALP = Alkaline Phosphatase, AST = Aspartate Aminotransferase, TP = Total Protein, ALB = Albumin, Glob = Globulin, A/G = Ratio of Albumin/Globulin, Chol = Cholesterol, Trig = Triglycerides, TBil = Total Bilirubin, Crea = Creatinine, GGT = Gamma Glutamyl Transpeptidase

Table S7. Reproductive Indices and Estrous Cycle Length in P1 Rats

Endpoint	Dose Level			
	0 ppm M/F	100 ppm M/F	300 ppm M/F	600 ppm F/800 ppm M
Number of Males	27	27	27	27
Number of Females	27	27	27	27
Male Mating Index	100.0 (27/27)	100.0 (27/27)	96.3 (26/27)	92.6 (25/27)
Female Mating Index	100.0 (27/27)	100.0 (27/27)	96.3 (26/27)	92.6 (25/27)
Male Conception Index	96.3 (26/27)	100.0 (27/27)	92.3 (24/26)	96.0 (24/25)
Female Conception Index	96.3 (26/27)	100.0 (27/27)	92.3 (24/26)	96.0 (24/25)
Male Fertility Index	96.3 (26/27)	100.0 (27/27)	88.9 (24/27)	88.9 (24/27)
Female Fertility Index	96.3 (26/27)	100.0 (27/27)	88.9 (24/27)	88.9 (24/27)
Gestation Index	100.0 (26/26)	100.0 (27/27)	100.0 (24/24)	100.0 (24/24)
Gestation Survival Index	98.6 (364/369)	99.5 (374/376)	98.5 (320/325)	99.7 (343/344)
Day 1 Survival Index	98.6 (359/364)	98.7 (369/374)	98.4 (315/320)	98.5 (338/343)
Day 4 Survival Index	98.6 (359/364)	98.4 (368/374)	98.4 (315/320)	97.7 (335/343)
Day 7 Survival Index	100 (264/264)	99.3 (265/267)	99.6 (238/239)	99.6 (235/236)
Day 14 Survival Index	99.6 (263/264)	99.3 (265/267)	99.6 (238/239)	99.2 (234/236)
Day 21 Survival Index	99.6 (263/264)	98.9 (264/267)	99.2 (237/239)	99.2 (234/236)
Sex Ratio on Day 1 - Male : Female	54 : 46	47 : 53	50 : 50	55 : 45
Time to Mating (days)	3.0 \pm 2.4	2.6 \pm 1.0	2.6 \pm 1.5	2.6 \pm 0.9
Gestation Length (days)	21.7 \pm 0.5	21.4 \pm 0.5	21.6 \pm 0.5	21.5 \pm 0.5
% Post Implantation Loss	10.03 \pm 9.91	12.17 \pm 15.39	13.63 \pm 12.34	8.63 \pm 12.14
Mean Estrus Cycle Length - Days	4.1 \pm 0.3 (39)	4.1 \pm 0.2 (39)	4.1 \pm 0.3 (39)	4.1 \pm 0.3 (39)

There were no statistical differences from control at $\alpha=0.05$.

Table S8. Reproductive Toxicity Endpoints

Endpoint	Study Type		
	2-Generation Reproduction Guideline (870.3800)	2,4-D EOGRS	EOGRS (OECD 443)
Dosing Duration	~36 weeks (2 generations)	~29 weeks (PND 139 in F1)	19-21 weeks (PND 98 in F1)
Anogenital Distance	Triggered (F2 Offspring)	Required (All F1 Offspring)	Required (All F1 Offspring)
Puberty Onset	1 pup/sex/litter	3-4 pups/sex/litter	3-4 pups/sex/litter
Estrous Cyclicity	20/dose (2 gens) (3 weeks)	20/dose (2 gens) (3-4 weeks)	20/dose (2 gens) (2 weeks + after VO)
Precoital Indices	20/sex/dose (2 gens)	20/sex/dose (1 gen unless triggered)	20/sex/dose (1 gen unless triggered)
Mating and Fertility Indices	20/sex/dose (2 gens)	20/sex/dose (1 gen unless triggered)	20/sex/dose (1 gen unless triggered)
Implantation Data	20/sex/dose (2 gens)	20/sex/dose (1 gen unless triggered)	20/sex/dose (1 gen unless triggered)
Postimplantation Loss	20/sex/dose (2 gens)	20/sex/dose (1 gen unless triggered)	20/sex/dose (1 gen unless triggered)
Gestation Length	20/sex/dose (2 gens)	20/sex/dose (1 gen unless triggered)	20/sex/dose (1 gen unless triggered)
Offspring Survival/Sex Ratio	20/sex/dose (2 gens)	20/sex/dose (1 gen unless triggered)	20/sex/dose (1 gen unless triggered)
Litter Size	20 litters (2 gens)	20 litters (1 gen unless triggered)	20 litters (1 gen unless triggered)
Epididymal Sperm Count	20/dose (control & high; 2 gens)	20/dose (control & high; 2 gens)	20/dose (control & high; 2 gens)
Testicular Spermatid Count	20/dose (control & high; 2 gens)	20/dose (control & high; 2 gens)	Not required
Sperm Morphology	20/dose (control & high; 2 gens)	20/dose (control & high; 2 gens)	20/dose (control & high; 2 gens)
Sperm Motility	20/dose (2 gens)	20/dose (2 gens)	20/dose (2 gens)
Ovarian Follicle Counts	F1 Adult Females (subset)	F1 Adult Females (15/dose - control & high; Set 3)	F1 Adult Females (F1 Cohort 1A)
Nipple Retention (males)	Not Required	All F1 Offspring	All F1 Offspring
Reproductive Organ Wts	20/sex/dose (2 gens)	20/sex/dose (P and F1 Set 3); 10/sex/dose (F1 Set 1A)	20/sex/dose (P); 40/sex/dose (F1 Cohorts 1A and 1B)
Reproductive Organ Histopath	20/sex/dose (2 gens)	10-12/sex/dose (P and F1 Set 1A); 20/sex/dose (F1 Set 3)	20/sex/dose (P and F1 Cohort 1A; examine F1 Cohort 1B if needed)

Bold text indicates differences in the 2-generation reproduction guideline (870.3800) and the 2,4-D EOGRS study compared with the EOGRS (OECD 443) test guideline.

Table S9. Developmental Neurotoxicity Endpoints

DNT Endpoint	Study Type		
	Developmental Neurotoxicity Guideline (OECD 426)	2,4-D EOGRS	EOGRS (OECD 443)
Dosing Duration for Offspring Examined for DNT Endpoints	GD 6 - LD 21	Continuous exposure to F1 across all life stages (including <i>in utero</i> and postnatal)	Continuous exposure to F1 across all life stages (including <i>in utero</i> and postnatal)
Detailed Clinical Observations (DCOs) in Parental generation	≥10 dams/dose (2-gestation; 2-lactation)	All P1 Adults (pre-exposure and 1X during pre-mating period)	All P1 Adults (weekly)
F1 DCOs	1 pup/sex/litter (20/sex/dose)	Not required (collected clinical observations on all F1 offspring)	Weekly in F1 offspring post-weaning
Puberty Onset	1 pup/sex/litter (20/sex/dose)	3-4 pups/sex/litter	3-4 pups/sex/litter
Neuropath/Brain Morphometry (PND 22)	1 pup/litter (10/sex/dose)	Neuropathology only (12/sex/dose)	Neuropathology only (10/sex/dose)
Behavioral Ontogeny (Surface Righting)	1 pup/sex/litter (20/sex/dose)	Not required	Not required
Motor Activity (PND 13, 17, 21)	1 pup/sex/litter (20/sex/dose)	Not required	Not required
Motor Activity (Adult)	1 pup/sex/litter (20/sex/dose)	10/sex/dose	10/sex/dose
Motor and Sensory Function (Adolescent) ¹	1 pup/sex/litter (20/sex/dose)	Not required	Not required
Motor and Sensory Function (Adult) ¹	1 pup/sex/litter (20/sex/dose)	10/sex/dose	10/sex/dose
Learning and Memory (Adolescent)	1 pup/litter (10/sex/dose) ²	Not required	Not required
Learning and Memory (Adult)	1 pup/litter (10/sex/dose) ²	Not required	Not required
Neuropath/Brain Morphometry (Adult)	1 pup/litter (10/sex/dose)	1 pup/litter (10/sex/dose)	1 pup/litter (10/sex/dose)

¹FOB: Auditory startle, landing footsplay, fore- and hindlimb grip performance, extensor thrust response, tail pinch, body temperature.

²Larger sample sizes could be considered e.g., up to 1 pup/sex/litter.

Table S10. Developmental Immunotoxicity Endpoints

Endpoint	Study Type		
	Immunotoxicity (870.7800)	2,4-D EOGRS	EOGRS (OECD 443)
Dosing Duration	≥28-days	Continuous exposure to Flacross all life stages (including <i>in utero</i> and postnatal)	Continuous exposure to Flacross all life stages (including <i>in utero</i> and postnatal)
T cell dependent antibody response	≥8 animals/dose; one sex only	10/sex/dose (Cohort 2A)	10/sex/dose (Cohort 3)
Positive control group (TDAR)	Required	Required	Required or justify if omitted
Enumeration of splenic lymphocyte subpopulations	Optional (consider if TDAR +)	Not required	10/sex/dose (Cohort 1A)
Splenic natural killer cell assay	Optional (consider if TDAR -)	10/sex/dose (Cohort 2B)	Not required
Spleen wts	≥8 animals/dose	10/sex/dose (weanlings, F1 Sets 1A, 2A, 2B) ; 20/sex/dose (2 gens)	10/sex/dose (weanlings); 20/sex/dose (2 gens)
Thymus wts	≥8 animals/dose	10/sex/dose (F1 Sets 1A and 2A); 20/sex/dose (F1 Set 3)	10/sex/dose (weanlings); 20/sex/dose (2 gens)
Lymph node wts	Not required	Not required	10/sex/dose (Cohort 1A)
Spleen histopathology	Not required	10-12/sex/dose (P1, F1 Sets 1A and 3, weanlings)	20/sex/dose (2 gens); 10/sex/dose (weanlings - optional);
Thymus histopathology	Not required	10/sex/dose (F1 Set 1A)	20/sex/dose (2 gens); 10/sex/dose (weanlings - optional);
Lymph node, bone marrow histopath	Not required	10/sex/dose (F1 Set 1A)	10/sex/dose (Cohort 1A)

Table S11. Comparative Thyroid Endpoints

Endpoint	Study Type		
	Comparative Thyroid Guidance	2,4-D EOGRTS	EOGRTS (OECD 443)
Dosing Duration	GD 6 – LD 21	Continuous exposure to P1 and F1 sets across all life stages	Continuous exposure to P1 and F1 sets across all life stages
Pregnant Dam Thyroid Evaluation ¹	10 dams/dose	10-12/dose (satellite dams)	Not Required
Fetal Thyroid Evaluation ¹	10 litters/dose	Not Required	Not Required
Cull Pup Thyroid Evaluation ¹	20 litters/dose	8-10/dose (thyroid hormones only)	Thyroid hormones optional ²
Adult/Lactating Dam Thyroid Evaluation ¹	20 dams/dose	P ♂ & ♀ (weights – 20/sex/dose; histopath - 10/sex/dose)	P ♂ & ♀ (weights and histopathology – 20/sex/dose;)
Weanling Thyroid Evaluation ¹	20 litters/dose (1♂ & 1♀/litter)	Unselected F ₁ offspring (PND 22; 10/sex/dose)	Unselected F ₁ offspring (PND 21-22; 10/sex/dose)
Adult F ₁ Thyroid Evaluation ¹	Not Required	Set 1a F ₁ offspring (10/sex/dose)	Set 1a F ₁ offspring (~PND 90; hormones - 10/sex/dose; weights and histopathology – 20/sex/dose)

¹ Thyroid evaluation includes thyroid hormones (TSH, T4 and sometimes T3), thyroid weight and thyroid histopathology unless otherwise specified.